VOL. XXI FEBRUARY, 1937 No. 1

THE JOURNAL OF APPLIED PSYCHOLOGY

JAMES P. PORTER, *Editor*
OHIO UNIVERSITY ATHENS, OHIO

With a Board of Cooperating Editors
and

PAUL S. ACHILLES HENRY C. LINK
Associate Editors Representing
THE PSYCHOLOGICAL CORPORATION

SOME INTERESTING CONTRIBUTIONS IN THIS ISSUE:

HENRY C. LINK. How Many Interviews Are Necessary for Results of a Certain Accuracy

DOROTHY M. ANDREW. An Analysis of the Minnesota Vocational Test for Clerical Workers

RUDOLF PINTNER, G. FORLANO AND H. FREEDMAN. Personality and Attitudinal Similarity Among Classroom Friends

PERCIVAL M. SYMONDS. Problems Falling within the Scope of Consulting Psychology

PUBLISHED BI-MONTHLY BY THE EDITOR
OHIO UNIVERSITY, ATHENS, OHIO.

PRINTED BY THE SCIENCE PRESS PRINTING COMPANY, LANCASTER, PA.

Entered as second-class matter at the post office at Athens, Ohio, under the act of March 3, 1879. Additional entry May 7, 1936, at the post office at Lancaster, Pa. Copyright 1936 by James P. Porter.

The Journal of Applied Psychology

CO-OPERATING EDITORS

W. P. Bogr, Indiana University, W. V. Bingham, Personnel Research Federation, New York; M. E. Broom, State Teachers College, San Diego, Calif.; H. S. Burnz, Ohio State University; Robert H. Gault, Northwestern University; William Healy, Baker Foundation, Juvenile Court, Boston; H. L. Hollingworth, Columbia University; J. E. W. Wallin, Atlantic University; Herbert S. Langfeld, Princeton University; James H. Leuba, Bryn Mawr College; Henry M. McComas, Johns Hopkins University; Walter D. Scott, Northwestern University; C. E. Seashore, University of Iowa; Daniel Starch, New York; Edward K. Strong, Jr., Leland Stanford Jr. University; L. M. Terman, Leland Stanford Jr. University; E. L. Thorndike, Columbia University; Joseph Jastrow, New School of Social Research, New York; G. M. Whipple, Danvers, Mass.; Helen Thompson Woolley, Teachers College, Columbia University.

The scope of The Journal of Applied Psychology will include the following: (a) The application of psychology to vocational activities, such as law, art, public speaking, industrial and commercial work, and problems of business appeals. (b) Studies of individual mentalities, such as types of character, special talents, genius, and individual differences, including the problems of mental diagnosis and vocational prognosis. (c) The influence of general environment conditions, such as climate, weather, humidity, temperature; also such conditions as nutrition, fatigue, etc. (d) The psychology of everyday activities, such as reading, writing, speaking, singing, playing games or playing musical instruments, sports, etc.

Communications, manuscripts, reviews, summaries, notices, books and offprints for review should be addressed to Dr. James P. Porter, Ohio University, Athens, Ohio. Contributors are requested to follow the conventional method of citation employed by the Journal; in every case the year, volume and pages should be specified.

The Journal of Applied Psychology is published bi-monthly. The subscription price is $4.00 a year. Single numbers $1.25 to $2.00, depending upon the number of pages. Remittances and business communications should be addressed to the Editor, James P. Porter, Ohio University, Athens, Ohio.

Vol. XXI FEBRUARY, 1937 No. 1

THE JOURNAL OF APPLIED PSYCHOLOGY

Edited by
JAMES P. PORTER, Ohio University
Athens, Ohio

TABLE OF CONTENTS

HENRY C. LINK. How Many Interviews Are Necessary for Results of a Certain Accuracy 1

DOROTHY M. ANDREW. An Analysis of the Minnesota Vocational Test for Clerical Workers 18

RUDOLF PINTNER, G. FORLANO AND H. FREEDMAN. Personality and Attitudinal Similarity among Classroom Friends 48

PERCIVAL M. SYMONDS. Problems Falling within the Scope of Consulting Psychology 66

MARY I. ELWOOD, EDWARD M. L. BURCHARD AND FLORENCE M. TEAGARDEN. An Evaluation of the Kent Oral Emergency Test 75

LAVERNE JOHNSON AND A. R. LAUER. A Study of the Effect of Induced Manual Handicaps on Automotive Performance in Relation to Reaction Time 85

ALBERT B. BLANKENSHIP AND HOWARD R. TAYLOR. Psychological Effects of Changing the Form of a Trade Name 94

F. H. FINCH AND M. E. ODOROFF. The Reliability of Certain Group Intelligence Tests 102

EMMETT L. SCHOTT. IQ Changes in Foster Home Children 107

WILLIAM C. F. KRUEGER. The Influence of Amount Limits and Time Limits upon the Rate of Work 113

RICHARD S. HARTER. The Effects of Training upon the Belief in Certain Popular Misconceptions 119

News and Notes 130

Book Reviews 132

New Books and Pamphlets 137

HOW MANY INTERVIEWS ARE NECESSARY FOR RESULTS OF A CERTAIN ACCURACY[1]

(In one time studies; In periodic studies; In sample product tests; In advertising tests.)

HENRY C. LINK
Psychological Corporation

THE problem of accuracy, as it will be discussed here, concerns only the laws of chance. It has nothing to do with the reliability of the raw data or the nature of the questions asked, or the significance of the facts obtained. For example, in one of our experimental studies we asked the same question of 4,000 people, but in the case of 2,000 of them, *one word* had been added to the question. The percentage who answered *yes* was:

Question 1—69% } Samples of 2,000 were comparable in
Question 2—81% } every respect.

If the results had been 74 per cent and 78 per cent, our conclusion would have been that the difference was only a chance variation, and that the two questions were probably the same. However, the difference of 12 per cent was more than a chance difference, and considerably more than twice the standard deviation. Therefore the questions were different, and the reliability of either question was in doubt.

[1] This paper is based upon an empirical demonstration of the applicability of certain statistical formulae to market research data. It was prepared for, and has been used by, business executives. Its purpose, therefore, was to present the practical application with a minimum of theory and technical detail. The tests of applicability were carried out much more thoroughly than is herein described. These tests were found to substantiate the applicability of the formulae to this type of questionnaire data. The study was planned and conducted under the direction of I. Lorge.

This illustration makes the point that the laws of chance and the problem of statistical accuracy have a direct bearing on the question of the *reliability* of raw data, methods, etc. So that while we are discussing here only that phase of accuracy based on the laws of chance, we are at the same time describing techniques which are useful in testing the reliability of our raw data, and more obviously, the extent to which we can draw conclusions from data believed to be reliable.

For example, suppose in two successive studies, we get the following results:

	JAN.	FEB.
Question A	15%	19.7%
Question B	15%	16.4%

Do these results indicate an increase from January to February? The results for question A seem to indicate a decisive increase, and yet, depending on the size of the sample in each case, the smaller numerical increase for Question B may represent an actual increase much more decisively than is shown by Question A. Here the criterion of accuracy *based on the size of the samples* is the one immediate requirement for a reliable conclusion about the results of these studies.

EXISTING FORMULAE

For answering such questions on accuracy in relation to the size of samples, we have over two hundred years experience and experiments, in highly developed and usable form, available. Beginning with the laws of probability in dice throwing and the Bernoulli distribution, and culminating in extensive developments by Karl Pearson, R. A. Fisher, and many others, statistical methods, formulae and concepts have been extensively developed.

While some of these statistical theories and formulae have been used in connection with market research data, their proper application has not yet received the extensive and careful study found in other fields of research. In June and July we made an especially intensive study of those statistical formulae with

the aid of psychologists who have specialized in behavior statistics. As a result, we are enabled not only to use these methods with greater confidence, but are also able to present much more clearly their practical significance as applied to problems of various kinds. For example, instead of concluding as has sometimes been done in the past, that a sample of 1,000 to 2,000 is all we need for practical purposes, we now conclude that *under certain circumstances, 200 is an adequate sample, while in other circumstances, a sample of 10,000 is not sufficiently accurate.* And we can state the reasons for this conclusion in strictly mathematical terms, as shown later.

The formulae in general use are given in all standard textbooks on statistics, among which may be mentioned Yule, Mills, T. L. Kelley, etc. The best known or most widely used formula is that for sigma or the standard deviation, namely

$$\text{Sigma} = \sqrt{\frac{pq}{n}}$$

From this are derived the tables showing sigma for samples and results of different size, such as the tables distributed by Professor Theodore H. Brown, Harvard Graduate School of Business Administration, and the results and tables given below. In the above equation, p is the proportion of desired answers and q the proportion of all other answers (*i.e.*, $p + q = 100\%$) and n is the total number of cases or interviews used. For a sample of 1,000 people, 50 per cent of whom say they use a certain product, this formula would show a standard deviation of 1.58 per cent. But according to the formula, also, in 32 cases out of 100 (that is if we had 100 samples of 1,000 people) the actual deviation would be more than the standard deviation and even as much as two times the standard deviation. In short, our 50 per cent answer might just as well have been anywhere between 53.16 per cent or 46.84 per cent. That is, in our next sample of 1,000 people, 53 per cent might have said they used that product, and in still another sample of 1,000 it might have been as low as 45 per cent, *even though all three samples may have been taken at the same time, in the*

same localities, by the same interviewers, and by exactly the same methods.

The standard deviation is the *unit* for measuring the probable accuracy of a single result and the significance of differences between samples. However, this unit may be used in a number of different ways, as follows:

1. *The probable error.* This is 67.4 per cent of the standard deviation. This figure gives the range within which the chances are 50–50 that a given result from a sample of given size is an accurate estimate for the population sampled. For example, 1,000 cases, 50 per cent of whom answer a certain way, the standard deviation (sigma) is 1.58 per cent, $\frac{2}{3}$ of this or the probable error is 1.055 per cent. The chances are fifty out of a hundred that this answer of 50 per cent is correct within plus or minus 1.055 per cent—that is, either 48.95 per cent or 51.05 per cent. However, in fifty cases out of a hundred it may be out of the way by larger sums, specifically as follows:

2. *One Standard Deviation.* In thirty-two out of a hundred cases the range of error will be at least one whole sigma or more, in some cases two sigmas, in a few cases between two and three. That is, in thirty-two out of one hundred samples or 1,000, the answer, 50 per cent, might lie either between 53.16 per cent or 46.84 per cent (2 sigmas) or 54.74 per cent and 45.26 per cent (3 sigmas).

3. *Two Standard Deviations.* In 95 out of one hundred cases, the result or answer will be correct within two sigmas or less.

4. *Three Standard Deviations.* In only five out of a hundred times will the error of the sample exceed 2 sigmas. In only three samples out of 1,000 will the error exceed three sigmas. Therefore, the measure of accuracy of three sigmas represents an occurrence so rare that for all practical purposes it can be ignored. The most useful criterion is two sigmas, because the chances are 95 in a hundred that the error will be less than two sigmas; or even 1 sigma, because the chances are

68 in a hundred that the error will be less than plus or minus 1 sigma. The researches of R. Fisher confirm the practicability of the two sigma limit. In some industrial problems (sampling inspection) three sigmas have been demonstrated as the practical limit, but the empirical tests of market research data conducted by The Psychological Corporation show this limit as a costly and unnecessary restriction.

AN EMPIRICAL ILLUSTRATION

How these well known and tried principles work out in actual practice was beautifully illustrated by an empirical study made by The Psychological Corporation on a sample of 5165 questionnaires. The results of this study will give a much clearer and more practical understanding of the errors due to sizes of samples than any abstract discussion can.

A STUDY OF 5165 INTERVIEWS

A survey made in April, 1934, in which 5165 homes were called on, was used as the basis for an empirical analysis of accuracy, *independent of any existing formula for computing accuracy*. All the questionnaires, after local tabulation, were coded, and punched on Hollerith cards. These cards were sorted at random into 51 sets of 100 cards, each set of 100 containing the same proportion of interviews from Eastern, Southern, Midwestern and Far Western Areas as there were in the total sample of 5165. These 51 sets of 100 were then tabulated for the toothpaste barometer[2] questions to determine the proportion of people using Brands A, B, C, etc., in each set of 100; for Brands A and E, the 51 sets of 100 showed the following results: (See Table, page 6.)

Now for Brand A, where 16.22 per cent of 5165 people answered a certain way, the sigma or standard deviation for a sample of 100 according to the formula given above, is 3.68 per cent, which gives us a range of 16.22 per cent plus or minus

[2] Jour. Appl. Psych., 18, 1, Feb., 1934, p. 16–26; *Harvard Business Review*, Jan., 1935, p. 193–204.

BRAND A		BRAND E	
No. of samples of 100	Percentage who Bought Brand A in Each Sample	No. of Samples of 100	Percentage who Bought Brand E in Each Sample
1	9	1	0
0	10	0	1
2	11	1	2
4	12	3	3
3	13	5	4
4	14	6	5
6	15	9	6
9	16	8	7
6	17	7	8
4	18	5	9
3	19	2	10
1	20	2	11
3	21	2	12
3	22	51	
1	23		
0	24		
1	25		
51			

Average of 51 samples of 100	16.39%	Average of 51 Samples of 100	6.63%
Actual average for 5165	16.22%	Actual average for 5165	6.56%

3.68 per cent, or from 19.90 to 12.54 per cent. It will be seen that thirty-five of our fifty-one samples of 100 or 68.6 per cent fall within this range. If we take twice the standard deviation, our range becomes 9.0 to 23.8 per cent. It will be seen that 49 of our 51 samples of 100, or 96 per cent, fall within this range. If we take three sigmas, our range becomes 5.3 to 27.5 per cent, and all of our samples of 100 fall within this range. If we take $\frac{2}{3}$ of 3.68 per cent, the standard deviation, namely, 2.46 per cent, 29 or 57 per cent of our 51 samples of 100 fall within this range.

All these results are strictly according to the Bernoulli or Binomial distribution and the associated probabilities described under paragraphs numbered 1, 2, 3, 4 above. In the case of Brand E, 48 or 94 per cent of our 51 samples fall within the range of twice the standard deviation.

DIFFERENT KINDS OF QUESTIONS

With questions of different kinds, and results of different sizes, the same results were found. Here are some of the questions tested by samples of 100:

1. What brand of tooth paste did you buy last?
2. How long have you been using this brand?
3. What product does Rudy Vallee advertise over the radio?
4. Is there any dentrifice you have in mind trying soon?

In short, *no matter what kind of question, or what the size of the result in per cent of answers, the Bernoulli distribution is the same, and the measure of accuracy is the same.* This should be true, theoretically, but our study proved it to be so empirically for results as low as .1 per cent.

PROOF BY SAMPLES OF 500

Continuing with the process of proof used in this study, the first, sixth, eleventh, etc., sets of 100 were combined to form ten sets of 500 which showed the following variations:

BRAND	SETS OF 500										RANGE BASED ON TWICE THE STANDARD DEVIATION	
	1	2	3	4	5	6	7	8	9	10		
A	15.0	17.8	15.8	16.8	17.4	17.2	16.2	17.8	14.8	15.8	13.08	19.70
B	8.6	10.8	11.2	12.4	8.0	11.4	10.0	10.6	10.0	11.8	7.79	13.27
C	5.2	7.4	5.2	6.2	8.6	6.8	8.0	7.8	6.2	7.0	4.59	9.11
D	10.0	8.0	9.4	8.6	7.6	10.8	7.0	7.4	6.6	8.0	5.90	11.86
E	8.6	5.0	5.4	6.2	7.2	5.4	6.8	6.0	7.6	8.6	4.34	8.78

In this case, none of the 50 results falls outside the range of twice the standard deviation. Furthermore, 66 per cent of the cases vary by no more than one sigma, which is in accord with the Bernoulli formula. These results are only a small part of the data used in this study and are reproduced here as an example.

The close agreement between the expected variation for samples of 500 (which is, of course, less than the variation by 100's) and the actual results above, means that the formula for

the Binomial distribution is directly applicable to the total sample of 5165 and describes the limits within which the proportion of users of each brand is correct for the total population sampled. Furthermore, this knowledge enables us to say that a certain range of accuracy will hold with a definite probability that it is correct. For example Brands A to E of dentifrices showed the following results for the total sample of 5165:

BRAND	PERCENTAGE OF USERS	68 CHANCES IN A 100 THAT THIS PERCENTAGE IS CORRECT WITHIN PLUS OR MINUS	19 CHANCES IN 20 THAT THIS PERCENTAGE IS CORRECT WITHIN PLUS OR MINUS	997 CHANCES IN 1000 THAT THIS PERCENTAGE IS CORRECT WITHIN PLUS OR MINUS
A	16.22	.51%	1.03%	1.54%
B	10.53	.43	.85	1.28
C	6.85	.35	.70	1.06
D	8.38	.39	.77	1.16
E	6.56	.35	.69	1.03

Similar results were obtained with other brands of toothpastes including several of the less frequently used brands and also with the answer to a variety of other questions. These results were further checked by rigorous statistical methods which corroborated the analysis above.

LAW OF DIMINISHING RETURNS

To double the accuracy of the results, the number of samples must be quadrupled. This conclusion is a corollary of the existing formulae already described. Our study, by its series of samples of 100, 500, and 1000, confirms this fact empirically. This fact has an important practical bearing, especially in relation to costs. Of course no amount of poorly gathered, inaccurate or biased data has any reliable value. But where data is reliably obtained, economical limits for the desired accuracy in the results can be set up. The degrees of accuracy resulting from an increasing number of interviews, in a case where 10 per cent of the people buy a certain brand, are shown on page 9. Obviously, increasing the number of interviews four times gives only twice the accuracy of the smaller sample, *e.g.*, compare

NUMBER OF INTERVIEWS	STANDARD DEVIATION	MINIMUM CHANGE NECESSARY TO INDICATE A TREND BASED ON 2 SIGMAS
1000	.95%	1.90%
2000	.67	1.34
3000	.55	1.10
4000	.47	.95
5000	.42	.85
7500	.35	.69
10000	.30	.60
15000	.25	.49
20000	.21	.42

1000 and 4000, or 5000 and 20,000 interviews above. Diminishing returns begin to be more and more pronounced after 4000 or 5000 interviews. For example, the permissible variation of .85 per cent for 5000 interviews is reduced only to .60 with 10,000 interviews, and to .42 with 20,000 interviews, an increase in accuracy of only 43 hundredths of one per cent through the addition of 15,000 interviews.

It should be pointed out, however, that where less or more than 10 per cent of the people interviewed are found to buy a certain brand, the range of accuracy varies accordingly. Suppose, in a study of 5000 homes, we find various brands used by various percentages of the homes, as shown below:

BRANDS	PER CENT OF BUYERS	STANDARD DEVIATION	MINIMUM DEVIATION TO INDICATE A TREND
A	1%	.14%	.28%
B	5	.31	.62
C	10	.42	.85
D	15	.51	1.01
E	20	.57	1.13
F	25	.61	1.22
G	30	.65	1.30
H	40	.69	1.39
I	50	.71	1.41
J	60	.69	1.39

etc., diminishing in the same ratios. The larger the proportion of people using a brand, up to 50 per cent, the larger the standard deviation. A table showing the standard deviation both by

the number of interviews, and for the different results in percentage of answers, has been prepared from the Bernoulli distributions, as a rough guide to those responsible for research and is given on page 16.

TABLE FOR DETERMINING ACCURACY BY SIZE OF SAMPLE

This table, given below, is based entirely on one sigma or standard deviation, which represents a 68 per cent chance of accuracy within the range of plus or minus one sigma. In the great majority of market studies, this is as great a degree of accuracy as is required.

In studies requiring a higher degree of accuracy, *e.g.*, periodic tests to measure the effectiveness of an advertising theme, our Psychological Brand Barometers, periodic measures of radio program popularity, the measure of twice sigma should be used. This can be readily done from our table simply by quadrupling the number of interviews given opposite the sigma. For example—If a result of approximately 15 per cent is expected and it is desired to obtain a probable accuracy of 2.00 per cent, read in 15 per cent column 2.06 per cent for 300 cases and multiply 300 by 4, 1200 interviews will give the desired accuracy.

HOW TO USE THE TABLE

The table can be used in two ways, (a) to determine the accuracy of a result already at hand; (b) to discover the number of interviews desirable for a result of a certain accuracy.

(a) If in a study based on 1000 cases, you wish to know the accuracy of a certain result, let us assume 20 per cent, you look down the column headed 20 per cent till you come opposite 1000. The figure at this point is 1.26. This is 1 sigma or standard deviation and means that your result will be plus or minus 1.26 or between 21.26 and 18.74 in 68 out of 100 cases. If, in a previous study, your result to the same question was 19 per cent, it then follows that there is no reliable difference between the results of your previous and present study.

To be still more certain of the range of accuracy of your result, you should multiply sigma by 2, giving 2.52 as the variation, plus or minus, or a range of 22.52 to 17.48 within which the chances of certainty are 95 in 100.

(b) However, if you wish to plan a study, you first decide the range of accuracy that will answer your problem. Suppose two results you expect, one between 45 per cent and 50 per cent, the other between 25 per cent and 30 per cent, will be sufficiently significant. A plus or minus sigma of 2.5 per cent will therefore not be too great. Look down the column headed by the largest result, 50, till you reach a sigma of 2.5. Read left from this sigma, to the column headed *size of sample,* and you will see 400 as the required sample. The chances are 68 in 100, that your results will fall within this desired range of accuracy. If this is not enough, and you desire an even greater certainty, such as is given by twice sigma, then multiply the sample by 4, giving 1600 interviews. Then, 95 times out of 100, your results will be accurate within plus or minus 2.5 for an answer of about 50 per cent.

CONCRETE EXAMPLES

Example 1. Suppose you are testing two different advertising themes, by any method whatsoever. If each theme was preferred by 50 per cent of the people tested or each brought the same number of coupons or sales there would be no difference. A result of 53 per cent and 47 per cent, or 30 per cent and 25 per cent with 45 per cent doubtful, would mean very little difference unless the sample were larger than 400 cases. Suppose, for this test, you considered 35 per cent and 20 per cent for two themes as showing the smallest difference which would clearly indicate the superiority of one of the themes. About 300 people would be necessary. We arrive at that conclusion in this way: A standard deviation of approximately 2.5 per cent for each result, or a total range of error for each result of 5 per cent would still leave a minimum actual difference of 5 per cent between two results of 35 per cent and 20 per

cent. For a standard deviation of 2.5 per cent, about 300 people are necessary (if each person is tested on both ads; or two sets of 300 are necessary if each person is tested on only one of the two ads).

If, in the first 200 people tested, the results are found to be 58 per cent for theme A, 22 per cent for theme B, and 20 per cent doubtful, then no further tests are necessary. Sigma for 58 per cent based on 200 cases is 3.46 per cent and for 22 per cent, 2.83 per cent, so that, at the extreme, in further samples the results might have been: Theme A—54.5 per cent, Theme B—25 per cent, a difference obviously decisive.

Example 2. We were recently requested to estimate on a product sampling job, in which people were to be asked to express their preference for one of two product samples. The estimate called for 1000 people, 2000, and 3000. We estimated on all three, adding the following:

"Although our estimates are based on 1000, 2000, and 3000 sets of samples, we believe that a carefully worked sample of 1000 is more than adequate, for the following reasons:

The minimum difference between preferences based on 1000 cases to be significant of a real difference in the two samples could be 46.6 and 53.4.
On the basis of 2000, it could be 47.6 and 52.4.
On the basis of 3000, it could be 48 and 52.

In short, the minimum difference between the per cent of people preferring number *one* and number *two*, to show a significant trend, is reduced only 2.8 per cents when 3000 instead of 1000 cases are taken."

"If, on the basis of 1000 cases, 38 per cent prefer product sample *one* and 62 per cent prefer product sample *two* then we have a difference so large that there can be no question of its significance, whether we take 3000 cases or 30,000. In other words, the greater the difference between product samples *one* and *two*, the less need there is for taking more than 1000 cases. Whereas, the smaller the difference between *one* and *two*, the less worth while it becomes to increase the number of people trying out these product samples."

If, in this study, we had started with 500 people, and found 55 per cent preferring sample 1, and 30 per cent preferring sample 2, we could have stopped at that point, because allowing

for the inaccuracies in a sample of 500, our results at the worst, would still have been 52.8 per cent and 32.05 per cent, a significant difference (sigma for (1) 2.21 per cent, for (2) 2.05 per cent, on 500 people, as per table). Or, if after 500, our results were 43 per cent for A and 30 per cent for B, we would not have needed to go further since, allowing for a sigma of about 2.19 per cent, it meant that there was practically no difference in the preferences of people for either of the two samples and that further samples would be as likely to show a smaller difference as a larger one.

Here is a case in which the cost of samples, distribution, and follow-up calls are such that possibly thousands of dollars can be saved through the proper use of our table.

Example 3. Suppose one wishes to measure the progress of radio programs and over a period of four months, in monthly studies, the results are as follows:

	FIRST MONTH	SECOND MONTH	THIRD MONTH	FOURTH MONTH
Per cent Listeners	5.6	8.9	5.5	6.8

Range of accuracy, based on $2 \times$ sigma, for above per cents.

	% %	% %	% %	% %
For samples				
of 200	8.9 to 2.3	12.7 to 5.1	8.8 to 2.2	10.1 to 3.5
of 500	7.8 to 3.4	11.3 to 6.5	7.7 to 3.3	9.0 to 4.6
of 1000	7.1 to 4.1	10.6 to 7.2	7.0 to 4.0	8.3 to 5.3

The sponsor who is anxiously awaiting his second report is delighted with the increase shown, but if the results, 5.6 per cent and 8.5 per cent have been based upon 200 or 500 cases, they are not at all to be relied on. They might as well have fallen anywhere in the wide range shown above, for the samples of various sizes. Even after four months, he can barely conclude, if he uses samples of 1000, that there has been a slight increase in the popularity of his program over the four months' period.

However, if based on samples of 4000 interviews, the results should turn out—

	FIRST MONTH		SECOND MONTH		THIRD MONTH		FOURTH MONTH	
	5.1%		6.3%		7.8%		7.4%	
	%	%	%	%	%	%	%	%
2 sigma range	5.85 to 4.35		7.05 to 5.55		8.65 to 6.95		8.25 to 6.55	

obviously, the second and third figures represent a reliable increase, and the fourth does not necessarily represent a decrease.

For programs which are showing a very rapid increase in popularity, such as 2 per cent, 5 per cent, 8 per cent, 11 per cent, 13 per cent, at two week intervals, samples of 2000 or even 1000 are sufficient. For programs showing a more gradual increase, samples of 4000, or even more, are necessary to measure small trends.

Example 4. In repeated studies, the more gradual the change being measured, the larger must be the size of the sample. Suppose a brand used by 15.5 per cent of people interviewed in January, is found used by 17 per cent in February. Unless the size of the sample is 4000 or over, this increase is not statistically reliable. If, in March, or two months later, 18.5 per cent are found to use this brand, and all three studies are based on 4000 people, the trend is reliable. The possible range of these three results, based on our table and using two sigmas, for a sample of 4000 would be:

	JANUARY	FEBRUARY	MARCH
Actual results and 2 sigma range of accuracy above and below.	from 16.6% 15.5% to 13.6%	from 18.2% 17% to 15.8%	from 19.7% 18.5% to 17.3%

However, in that case, the February study would have been unnecessary.

To be sure of an increase as shown by results of 15.5 per cent and 17 per cent, in one month, samples of 4,000 would be required. A difference of two sigmas between the per cents observed will be used in the Psychological Brand Barometer[3] reports as the minimum change necessary to indicate a trend for a brand, when based on two successive studies, of one sigma when based on three successive studies *if* the trend of all three is the same.

HOW TO READ THE TABLE

The table on page 16 shows the standard deviation for a series of percentages and sizes of sample. This table has been made brief for easier reading and to give quickly a reasonably accurate result for ordinary purposes. The column at the left gives the size of the samples. The headings above the remaining columns show the *result*, in per cents, the accuracy of which is to be determined. Reading down the *result* column to a point directly opposite the *sample* column gives the sigma or unit of accuracy. For example, the answer is 30 per cent, the sample is 1500, sigma is 1.18 per cent. Or a previous result was 25 per cent, based on 1000 cases, the sigma is 1.37. We now have a later result of 30 per cent, also based on 1000 cases, the sigma is 1.45. The 25 per cent might have been 26.37, and the 30 per cent less 1.45 be 28.55, there would still be a significant difference between the two results.

Note: Standard deviations given in the table to the nearest .01 of one per cent when multiplied by 2 are accurate within .01 of one per cent and when multiplied by three are correct within .015 of one per cent. Greater accuracy is rarely of value in practical market research problems. In rare cases where results show a difference very close to twice the sigma, add .005 per cent to the sigma in the table to determine the significance of the results.

For percentages not shown in the column headings use the column nearest the percentage desired.

[3] *Op. cit.*

Standard Deviation Given in Per Cent and Hundredths of a Per Cent

SIZE OF SAMPLE	EXPECTED OR OBSERVED VALUE FROM A SAMPLE						
	50%	40% or 60%	30% or 70%	25% or 75%	20% or 80%	17.5% or 82.5%	15% or 85%
100	5.00	4.90	4.58	4.33	4.00	3.80	3.57
200	3.54	3.46	3.24	3.06	2.83	2.69	2.52
300	2.89	2.83	2.65	2.50	2.31	2.19	2.06
400	2.50	2.45	2.29	2.17	2.00	1.90	1.79
500	2.24	2.19	2.05	1.94	1.79	1.70	1.60
750	1.83	1.79	1.67	1.58	1.46	1.39	1.30
1000	1.58	1.55	1.45	1.37	1.26	1.20	1.13
1500	1.29	1.26	1.18	1.12	1.03	.98	.92
2000	1.12	1.10	1.02	.97	.89	.85	.80
2500	1.00	.98	.92	.87	.80	.76	.71
3000	.91	.89	.84	.79	.73	.69	.65
4000	.79	.77	.72	.68	.63	.60	.56
5000	.71	.69	.65	.61	.57	.54	.51
6000	.65	.63	.59	.56	.52	.49	.46
7000	.60	.59	.55	.52	.48	.45	.43
8000	.56	.55	.51	.48	.45	.42	.40
9000	.53	.52	.48	.46	.42	.40	.38
10000	.50	.49	.46	.43	.40	.38	.36
15000	.41	.40	.37	.35	.33	.31	.29
20000	.35	.35	.32	.31	.28	.27	.25
30000	.29	.28	.26	.25	.23	.22	.21
40000	.25	.25	.23	.22	.20	.19	.18

	12.5% or 87.5%	10% or 90%	8% or 92%	6% or 94%	4% or 96%	2% or 98%	1% or 99%	.5% or 99.5%
100	3.31	3.00	2.71	2.37	1.96	1.40	.995	.705
200	2.34	2.12	1.92	1.68	1.39	.99	.70	.50
300	1.91	1.73	1.57	1.37	1.13	.81	.575	.41
400	1.65	1.50	1.36	1.19	.98	.70	.50	.35
500	1.48	1.34	1.21	1.06	.88	.63	.445	.315
750	1.21	1.10	.99	.87	.72	.51	.36	.26
1000	1.05	.95	.86	.75	.62	.44	.315	.22
1500	.85	.77	.70	.61	.51	.36	.26	.18
2000	.74	.67	.61	.53	.44	.31	.22	.16
2500	.66	.60	.54	.47	.39	.28	.20	.14
3000	.60	.55	.50	.43	.36	.26	.18	.13
4000	.52	.47	.43	.38	.31	.22	.16	.11
5000	.47	.42	.38	.34	.28	.20	.14	.10
6000	.43	.39	.35	.31	.25	.18	.13	.091
7000	.40	.36	.32	.28	.23	.17	.12	.084
8000	.37	.34	.30	.27	.22	.16	.11	.079
9000	.35	.32	.29	.25	.21	.15	.105	.074
10000	.33	.30	.27	.24	.20	.14	.10	.070
15000	.27	.24	.22	.19	.16	.11	.08	.058
20000	.23	.21	.19	.17	.14	.10	.07	.050
30000	.19	.17	.16	.14	.11	.08	.06	.041
40000	.17	.15	.14	.12	.10	.07	.05	.035

Notice that the sigma in any column varies in inverse proportion to the square root of the size of the sample. That is, for 50 per cent sigma is 5.0 per cent for 100 cases and .5 per cent for 10,000 cases, where the size of sample increases by 100, the sigma decreases by one tenth. This fact can be used to extend the size of sample figures by using the squares of numbers from 1 to 10 (4, 9, 16, 25, 36, 49, 64, 81, 100) in the following way. For the sigma of 15 per cent, 2700 cases, divide 2700 by 9 and read opposite 300 the sigma for 15 per cent, 2.06 per cent, then dividing this by 3 (the square root of 9) gives .69 per cent the sigma of 15 per cent for 2700 cases. Looking at the sigmas in the table for 2500 and 3000, .69 per cent is obviously correct for 2700 cases. This simple procedure is more accurate than trying to interpolate for 2700 cases between the sigmas for 2500 and 3000 cases.

AN ANALYSIS OF THE MINNESOTA VOCATIONAL TEST FOR CLERICAL WORKERS.* I

DOROTHY M. ANDREW
Pennsylvania College for Women

THE Minnesota Vocational Test for Clerical Workers has been used extensively at the University of Minnesota Employment Stabilization Research Institute in the study and guidance of workers and in the differentiation of groups of employed and unemployed workers in various occupations. It has been demonstrated (2) that the clerical test differentiates clerical workers from workers in the general population better than an academic ability test, and that it discriminates somewhat less successfully various classes of clerical workers, such as stenographers, from general clerical workers. Another investigator (6) also has reported significant differences in clerical test scores between men office clerks and garage mechanics, and between women office clerks and retail saleswomen. Other findings have shown the Minnesota Clerical Test to be not only a reliable and valid measuring instrument but one which seems to measure an aptitude relatively independent of clerical training and experience (2). To show, however, that a given aptitude test measures clerical

* The contents of this article were taken from the author's Ph.D. dissertation on file in the office of the Graduate School of the University of Minnesota. The author is grateful for the facilities made possible by the Employment Stabilization Research Institute for the additional standardization and study of the Minnesota Clerical Test. She gratefully acknowledges her indebtedness to her adviser, Professor Donald G. Paterson, for guidance during the course of this research. She wishes to express her appreciation to Dr. Charles Bird for offering many helpful suggestions which have aided in the preparation of the manuscript.

ability does not describe clerical ability adequately, nor does it differentiate clerical ability from other measurable abilities.

The aim of this study is not to produce further evidence for the reliability and validity of the test but to find out what abilities the clerical test is measuring, and what abilities it is independent of when applied to a homogeneous group of clerical workers. The problem, which is one of mental organization, can be studied most adequately through an analysis of the intercorrelation of tests devised to measure various abilities. It is important to devise an instrument which adequately measures an ability and it is equally as important to isolate, study, and differentiate that ability from other tested abilities. Progress in mental analysis depends largely upon the measurement and isolation of unique traits. Until such comprehensive studies are made and described, ability tests will aid theoretical psychology little in the understanding of man even though their practical applications may be numerous. Mental analysis studies also have many practical applications. Industrial and educational fields will be aided by analyses of the independence and interrelations of mental traits.

The first problem, then, is to discover what aptitudes the clerical test is measuring or what traits are closely related to the performance it samples. To do this, it will be necessary to select a battery of tests which measure simple abilities so that when the relationships are found it will be possible to single out and discuss the nature of the existing correlation.

In the study of the relation of the clerical test to other abilities, it is important to study also the relationship of the test to other measures of clerical ability. Previous results have shown that the Scott Company's File Clerk's Test and the Hoke Prognostic Test of Stenographic Ability are measuring in part the same abilities as the Minnesota Clerical Test (2). Since the aim of this study is to explain the nature of clerical ability, single tests of relatively simple abilities presumably measuring clerical aptitude will be studied in relation to the clerical test.

The first step, then, in the analysis of any ability is the examination of basic elements which constitute or are related to that ability. In this study, the basic elements will consist of a battery of tests of simple functions, such as dexterity, cancellation, and substitution, as well as of tests of simple arithmetic, spelling, observation and comparison, and number group checking.

A second problem is to determine if the clerical test is measuring a unique trait, that is, one which is relatively independent of other tested abilities. At present the relationship between mechanical aptitude, dexterity, academic ability tests and clerical ability are important considerations. If the clerical test is measuring an ability different from other aptitudes, the test when applied to a group of clerical workers or workers in other occupations must not show high degrees of relationship with other tested abilities. High degrees of relationship would mean that the tests were measuring approximately the same abilities but in a different guise. To warrant existence, each new test must show how it is measuring an ability better than previous tests, and whether it is measuring abilities which are relatively independent of those already being measured.

A final consideration to be considered in a succeeding publication, as revealed by factor analysis techniques, is the question of uniqueness and overlapping of abilities and the number of factors involved in the many tests. The intercorrelations for this analysis will be studied by the Thurstone "Simplified Multiple Factor Method" and by Spearman's technique of tetrad analysis. Factor analysis is considered because of its possible contribution to the understanding of the measured abilities and not for any theoretical considerations aimed to disprove or substantiate the techniques or the various theories of factor analysis.

EXPERIMENTAL PROCEDURE

All of the subjects for this study were employed or unemployed workers who were being tested by the occupational

analysis clinic of the Employment Stabilization Research Institute. The regular battery of tests used by the Institute was administered by the Institute Testing Staff to each subject, whereas the special battery of tests was given by the writer. All subjects were tested individually with both the Institute and the special test batteries. There were 340 workers tested for this study, 216 women and 124 men. Of the 340 workers, 155 women clerical workers were selected to serve as an experimental group.

The occupational classification of the subject was not always known at the time of testing since the occupational interview frequently was not taken until late in the day after the testing had been partially or fully completed. Consequently, the sample consisted of a fairly heterogeneous group with respect to age and occupation. When it became apparent that there were more women in the sample than men, and that the women were a more homogeneous group with respect to age and occupation, the last weeks of testing were restricted to women clerical workers.

The women selected for this study ranged from 17 to 29 years of age. All had had some experience in clerical work with the exception of four who had no work history but who seemed suited to clerical work according to their educational record and test scores. Of the 155 women workers, 139 were employed clerical workers. The others were unemployed.

From the files of the Institute, data were obtained concerning sex, age, usual occupation, occupational history, years of clerical experience, and years of formal schooling.

Among the Institute battery of tests (7) were measures of academic ability, clerical ability, manual, finger, and tweezer dexterities, mechanical ability, personality traits, and vocational interests. Personality and vocational interests scores were not considered important for the present study, and the data in the mechanical assembly test were too incomplete to be useful. All other test scores were used.

TESTS USED

(a) *Institute Battery*

Pressey Senior Classification and Verification:
These tests are measures of academic ability. In this experiment the average of the two test scores was used. The two Pressey Tests correlate about + .90 with each other, an indication that they are measuring approximately the same abilities (13).

Minnesota Vocational Test for Clerical Workers:[1]
This is a test of clerical aptitude consisting of two parts, a clerical number checking test, and a clerical name checking test.

O'Connor Finger Dexterity Test:
This is a test which measures speed and accuracy in eye-hand coordination. The subject is required to pick up three pins at a time and to place them in each of one hundred holes.

O'Connor Tweezer Dexterity Test:
The tweezer dexterity test measures the speed and accuracy with which the subject can manipulate a pair of tweezers in the picking up and placing of one pin in each of one hundred holes. This test requires a high degree of speed and accuracy in eye-hand coordination and steadiness.

Minnesota Manual Dexterity Test (Placing):
This test measures the speed with which the subject can place relatively large circular wooden blocks in holes in a board. "Success is dependent upon speed of gross hand and arm movements." (7)

Minnesota Spatial Relations Test:
This is a test of mechanical ability which is relatively free from the effect of mechanical experience. It is made up of four boards, each board containing 58 cut-outs. It is a

[1] The Minnesota Vocational Test for Clerical Workers is distributed by the Psychological Corporation, 522 Fifth Avenue, New York City. A complete discussion of the nature of and results obtained from the Minnesota Clerical Test is published in a bulletin of the Employment Stabilization Research Institute, entitled "Measured Characteristics of Clerical Workers," written by D. M. Andrew and D. G. Paterson. This bulletin is available through the University of Minnesota Press.

test of speed and accuracy in form perception, since it requires one to discriminate quickly and accurately various odd shapes and sizes.

In the final set-up, thirteen additional special tests were administered individually to each subject. Several criteria aided in the selection of tests for this special battery. The first was that the test must be one which has been used to measure clerical ability. Secondly, the test must measure a relatively simple ability which can be described. The third criterion was the plausibility of the test being related to the clerical test. Only tests were selected which seemed to have a possibility of being related to the clerical test and of measuring a simple ability having value in the analysis of clerical aptitude. A fourth requirement, that of the reliability of the special tests, was considered. When reliability coefficients for similar tests were reported in the literature, such reliabilities were used to lengthen and predict the reliabilities of the tests. If coefficients were not reported for similar tests, reliabilities were determined in a preliminary experiment.

(*b*) *Special Tests*

Feature Discrimination: (8)

Four tests of feature discrimination were selected from the Michigan Non-Verbal Series prepared by Edward B. Greene. These discrimination tests are designed to measure the ability to observe and compare by having the subject find the unique feature in each row of small faces. The tests consist of four parts: 2A, 3A, 4A, and 5A, with increasing complexity. In the first part, two faces are to be compared; in the second, three; and in the third and fourth parts, five and seven faces are to be compared in each row, respectively. These tests were included in the special battery since observation and comparison, which are the abilities supposedly measured by the feature discrimination tests, are also traits which may have a significant relationship to success in the clerical checking tests.

Number Cancellation:

The number cancellation test consists of two parts of equal length each with a time limit of one and one-quarter min-

utes. There are a total of 1500 numbers in the two parts or 750 in each half. The test is constructed so that there are four sevens to be cancelled in each line, giving a total of 100 sevens to be cancelled in each half. This test is based on the Woodworth-Wells tests of cancellation. Tests of cancellation are assumed to measure quickness in perception, speed of discrimination, and quickness of reaction. Cancellation tests, both number and letter, have been used in various studies of clerical ability. Since speed in perception and discrimination seem to be factors in clerical ability independent of academic ability (9) and seem to be measuring an ability lacking complexity, both letter and number cancellation were included in this study.

Letter Cancellation:

The letter cancellation test was constructed similarly to the number cancellation test. There were two parts, each half made up of 750 letters, among which were 100 a's, four in each line, to be cancelled.

Substitution Tests:

Two tests of substitution are included in the test battery, one of letter substitution and the other of number substitution. Each test is made up of 180 letters or numbers. The letters and numbers are repeated an equal number of times. In the letter substitution test, five letters are to be substituted for five letters and in the number test, five numbers are to be substituted for five numbers. Substitution tests have been called tests of quickness in learning. Most substitution tests, however, seem too simple and brief to be adequate measures of the complicated learning process, but at least they may give an indication of the subject's ability to adapt to a situation in which speed in learning a short code is an advantage. Substitution tests were included in this test battery for several reasons, one being that they have been used by many investigators in the study of clerical ability. (1) (15) (14) (10) (9). Other reasons for including measures of substitution were that such tests seemed to measure the ability to adapt quickly to a new situation, an ability which might be important in clerical tasks.

Verifying Arithmetic Test:

The verifying arithmetic test was designed to measure simple arithmetical ability. The short test consists of 40 prob-

lems in addition, multiplication, subtraction, and division. There are ten problems in each of the above mentioned four categories. All problems are simple enough so that they can be checked mentally. Since the problems are answered, the task of the subject is to place either a C on the line after each problem if the problem is correct, or to place a check on the same line if the answer is wrong. This method is preferred to the writing of the answer since maximum time is spent in calculating and not in writing. Eleven of the forty problems are wrong, some of which are in each of the four kinds of problems. An arithmetic test was included in this series of tests since the relationship between arithmetic and score in a clerical number checking test seems to be an important consideration. Many investigators have included arithmetic tests in their clerical battery. (10) (11) (4) (15).

Spelling Tests:

Two spelling tests also were given to the subjects. The first spelling test, which will be designated as Spelling I, was designed not so much to measure spelling as to measure the ability to perceive quickly and accurately slight misspellings in a group of easy words. These words were chosen from Ayres Spelling Scale and from the word lists which were passed by from 66 per cent of second grade to 58 per cent of fifth grade children. Words selected for the final list are so simple that errors made by an average adult should be made through inaccuracies in perception. The list of fifty words included twenty-six spelled incorrectly and twenty-four spelled correctly. The subject recorded his answer by checking under the column headed 'R' if the item was correct and under 'W' if the word was spelled incorrectly.

The second spelling test, called Spelling II, was definitely designed to measure spelling ability. It includes 175 words selected from the Sixteen Spelling Scales reported in the Teachers' College Series. The words were chosen from word lists which were passed by better than 50 per cent in eighth grade and others were passed by less than 50 per cent in ninth grade. About 56 per cent of the words are spelled incorrectly in the test which means that 99 are wrong and 76 are spelled correctly. The directions for taking the test are the same as those in the easier spelling test. It is apparent from the wide use of spelling tests in

the study of clerical ability (15) (3) (12) that the inclusion of such tests in the analysis of clerical ability is desirable.

Number Group Checking:

The number group checking test has been used extensively in the study of clerical ability. Primarily, it measures the ability to perceive quickly and accurately the presence or absence of two numbers in a group of numbers. In this test 200 groups of numbers are arranged in five columns. The subjects are to check each group that contains both a six and a nine. There are 75 such groups to be checked, fifteen in each column. This test is based on the Woodworth-Wells Number Group Checking Test which contains 168 items, seventy of which are to be checked because they include both a 'one' and a 'seven.' The number group checking test was included in the present investigation because it measures an ability which is apparently closely related to clerical number checking. Number Group Checking Tests have been used by investigators (10) (14) in the measurement of clerical ability.

Test in Adding:

This test requires the ability to follow a line of numbers and to underline any two consecutive numbers whose sum equals ten. The test consists of 1260 numbers which are divided into 20 lines with double spacing between every five lines. In each line there are ten adjacent numbers whose sum equals ten with the exception of three lines in which there are eleven such pairs. Burtt (5) believes that such a test is valuable in the prognosis of clerical ability since it demands, like clerical work, a high degree of concentration. Since the Minnesota Clerical Test, as well as other tests, require concentration, the relationship between this adding test and the other clerical tests is worth considering.

TESTING PROCEDURE

The order in which the Institute Battery was given varied with different subjects. The special battery of tests, however, always was given in a definite order and always followed the administration of the clerical test. This rule was adhered to so that no practice effect on the special battery could influence

clerical test scores. The first four tests in the special battery, the four feature discrimination tests, were given in order of increasing difficulty, 2A, 3A, 4A, and 5A. These tests were given first so that the subject would be alert and patient in the task of looking for the one difference in the rather large groups. Number and letter cancellation tests and letter and number substitution tests followed the discrimination tests in the above order. Ninth in the series was the verifying arithmetic test, which was followed by Spelling I and II, and these by number group checking and adding. To serve as a shock absorber, the easier spelling test was given before the more difficult one. The adding test was administered last since it was rather tiring to the eyes. The plan used in determining the above testing order was that tests which are similar should be given consecutively. All tests in the special battery were given by the time limit method.

EVALUATION OF THE TESTS

In a study of the value of any measuring instrument, records of the average score, variability, skewness of the distribution and reliability are important considerations. Table I gives the averages, variabilities, and skewness of the distribution for the variables.

The negative skewness for clerical experience is greater than 3.00 and, therefore, is statistically significant. It indicates that there is a considerable piling up of cases in the intervals of little clerical experience. Since the experimental group has been selected from among young clerical workers, it might be expected that most of the clerical workers would have had little experience. The average worker in this group has had four years' experience in clerical work; about one-sixth of the group have had nine or more years; whereas, about one-third have had less than two years' clerical experience.

The distribution for age is not significantly skewed. The average age for this group of women clerical workers is 23 years and the age range is from 17 to 29 years inclusive.

TABLE I

Skewness, Means, and Standard Deviations for the Test Scores and Other Variables

MEASURE	MEAN	STANDARD DEVIATION	SK.	$\sigma_{SK.}$	$\frac{SK.}{\sigma_{SK.}}$
Years of clerical experience	4.38	3.59	−1.51	.47	−3.23
Age	22.86	3.33	− .63	.43	−1.46
Years of formal schooling	11.99	1.38	.12	.06	2.00
Pressey education tests (average of the two)	69.30	14.01	2.94	1.73	1.70
Clerical: number checking	145.68	28.11	.68	3.68	.19
Clerical: name checking	149.80	28.51	5.62	3.79	1.49
Finger dexterity	3.85	.50	.14	.05	2.50
Tweezer dexterity	5.54	.83	.16	.09	1.78
Manual dexterity	227.71	20.19	−5.44	2.43	−2.24
Spatial relations	1179.24	260.52	64.25	27.88	2.30
Feature discrimination 2A	32.73	8.71	− .88	.99	− .09
Feature discrimination 3A	27.30	6.33	− .06	.67	− .09
Feature discrimination 4A	11.40	5.03	− .16	.63	− .25
Feature discrimination 5A	12.70	4.27	.15	.55	.27
Number cancellation	145.99	27.43	1.40	3.54	.40
Letter cancellation	148.06	28.35	−1.53	3.68	− .42
Letter substitution	115.48	26.76	−8.42	3.55	−2.37
Number substitution	139.61	23.08	−4.38	2.99	−1.47
Verifying arithmetic	25.95	7.54	−2.32	.99	−2.34
Spelling I	44.74	7.65	8.43	.83	10.11
Spelling II	91.55	35.08	6.91	4.29	1.61
Number group checking	144.84	31.17	− .31	4.03	− .08
Adding	101.68	28.48	2.10	3.54	.59

The average clerical worker in this group has had a twelfth grade education, but the range in education is from eighth grade through college graduation. The slight positive skewness indicates that there is a tendency for the piling up to be in the higher educational levels. In academic ability, the average worker in this group has a "B" rating[2] which is better than 75 per cent of women workers in the general population. There is a tendency for the greatest frequency of cases to be among the better scores in the Pressey Tests.

The mean in clerical number checking for this group of subjects is "B" grade, or a score equal to a percentile of 85 in com-

[2] General population norm letter grade limits are: Standard score 6.5 and above, A; 5.5 to 6.4, B; 4.5 to 5.4, C; 3.5 to 4.4, D; and 3.4 and below, E.

parison with women workers in general. Distributions for both the clerical number and name checking tests are normal. The average clerical worker in this group exceeds 85 per cent of women workers in the general population in clerical name checking, which is also equal to "B" grade ability.

The distribution for finger dexterity tends to be positively but not significantly skewed. For this group the mean is equal to a percentile of 65 or to "C" grade ability for workers in general. Likewise in tweezer dexterity, the average clerical worker in this group has a percentile rating of 60 or "C" grade for women workers in the general population. In manual dexterity, the average score is equal to the fiftieth percentile or "C" grade ability.

In the spatial relations test, the average for this group of women is also equal to "C" grade or the 65 percentile for women workers in general. The distribution tends to be skewed but not significantly so toward the higher end of the scale.

In concluding the discussion of the scores, for this group, on the Institute Battery of Tests, it is important to note that the average scores resemble closely those for employed women clerical workers tested in an earlier study. (2).

The range of scores on the feature discrimination tests was large, extending from near zero to maximum scores. All distributions are approximately normal.

On the cancellation tests, the average scores are about equal for number and letter cancelling. The number substitution test, however, has a higher average score than has letter substitution. Part of this gain may exist because a number code is more easily learned or because in the second part the subject is more adapted to substitution. The curves for cancellation are not skewed, whereas the distributions for the substitution tests are slightly but not significantly skewed, indicating a piling up in the lower scores.

The average clerical worker in this group scores 26 out of a possible 40 in the Verifying Arithmetic Test. The skewness

which is present is not significant but it does indicate that the scores tend to be at the lower end of the scale.

Because of the high average score, 48 out of 50, and the extreme positive skewness of the distribution, the Spelling I Test is not deemed satisfactory as a differentiating instrument for this group of clerical workers. In a preliminary study, the average score for a group of workers chosen at random was 24.5 on this Spelling Test. For a group of clerical workers, however, the test is too easy. Although the Spelling I Test is used in the intercorrelations, it is eliminated in the later calculations because of its skewed distribution of scores. In the Spelling II Test the average score is 92 out of a possible 175 and the distribution of scores is not significantly skewed.

Both the number group checking and the adding tests give normal distributions for this group of women workers. The average score in number group checking is 145 out of a possible 200 and, in adding, a score of 102 out of a possible 200.

Because of incompleteness of data the manual dexterity test could not be used in the factor analysis study. Manual dexterity, however, correlates +.36 with clerical number checking and +.29 with clerical name checking.

RELIABILITY OF THE TESTS

In studying interrelationships of abilities, the measuring instruments should have a high degree of internal consistency. Table II gives reliability coefficients for the tests used in this study.

The reliability of the Pressey Education Tests is sufficiently high for individual diagnosis. Calculated by comparing scores on the Senior Classification with scores on the Senior Verification Tests, the reliability was +.90 for women workers and +.91 for men workers. (13). In this study the average of the scores on the two tests was used, so we may expect the reliability for this group to be equally high.

The reliability of the finger dexterity test for a heterogeneous group of women workers is +.81, corrected, +.90. For

a homogeneous group of 200 clerical workers, 155 of which are included in this study, the reliability is +.74 or corrected, +.85. This slight decrease in reliability was expected because of the homogeneity of the group.

Reliabilities could not be directly ascertained for the tweezer dexterity test from the Institute data, since the one trial given yielded only one time score. Indirect evidence of its reliability is gained from a study of the relationship of the finger to the tweezer dexterity test. (13). For three random subgroups of women, the average intercorrelation is +.33. The average intercorrelation for the total sample is +.32. Difference ratios for the three groups are not significant. Darley (13) concluded that the "consistency with which this relationship is maintained again suggests but does not prove that the Tweezer Dexterity Test approximates in reliability the Finger Dexterity Test." Study of the scores of men workers indicate similar results.

The manual dexterity test has a reliability of +.84, or corrected, +.91 when the first and third trials are compared with the second and fourth trials for a group of 150 women workers. The reliability of the Spatial Relations Test is +.80 or corrected, +.89 when boards A and B are compared with boards C and D for 224 women workers.

The reliability of the Minnesota Clerical Test has been tested many times by various procedures, and with various groups. Both odd-even and test-retest reliability coefficients for various groups have indicated that the reliability of the test is sufficiently high for individual diagnosis. The reliabilities of +.86 and +.93, uncorrected, reported in Table II for women clerical workers are additional evidence for the reliability of the clerical number and name checking tests.

Darley concluded with respect to the Institute Battery of Tests that "the aptitude and ability tests used by the Research Institute meet the criterion of reliability for use in making group comparisons, and with the exception of the mechanical assembly boxes each test seems to be adequate for individual diagnosis." (13).

TABLE II

Reliabilities of Tests Used in This Study

TESTS	DESCRIPTION OF GROUP	ODD-EVEN		SPEARMAN-BROWN
Pressey Tests* (Classification vs. Verification)	131 women workers	.90 (Comparable Forms)	.02	.....
Finger dexterity* (first 50 vs. 2nd 50 holes)	215 women workers	.81	.02	.90
Finger dexterity	200 women clerical workers	.74	.02	.85
Tweezer dexterity*	See explanation	.....	.....	.....
Manual dexterity*	150 women workers	.84	.02	.91
Spatial relations*	224 women workers	.80	.02	.89
Clerical number checking	147 women clerical workers	.86	.01	.92
Clerical name checking	147 " " "	.93	.01	.97
Feature discrimination 2A	200 " " "	.95	.004	.97
Feature discrimination 3A	200 " " "	.94	.01	.97
Feature discrimination 4A	200 " " "	.94	.01	.97
Feature discrimination 5A	200 " " "	.88	.01	.94
Number cancellation	200 " " "	.70	.03	.82
Letter cancellation	200 " " "	.81	.02	.89
Letter substitution	200 " " "	.98	.001	.99
Number substitution	200 " " "	.99	.001	.99
Verifying arithmetic	200 " " "	.93	.01	.96
Spelling I	200 " " "	.89	.01	.94
Spelling II	200 " " "	.92	.01	.96
Number group checking	200 " " "	.86	.01	.92
Adding	200 " " "	.98	.002	.99

* Reliabilities for the academic ability, finger, tweezer, manual dexterity, and spatial relations tests (starred) are taken from "Research Studies in Individual Diagnosis," Editor, Donald G. Paterson, Bulletin of the Employment Stabilization Research Institute, pp. 44 and 46.

All tests in the special battery are sufficiently reliable for individual prediction with the possible exception of the number cancellation test which is reliable enough for group prediction.

THE RELATION BETWEEN TEST SCORES AND AGE, CLERICAL EXPERIENCE, AND FORMAL EDUCATION

Since age has been controlled by eliminating all clerical workers over 29 years of age, a decrement in test scores due to age should not be expected. The only correlations between age and the other variables which border on significance are those between age and Spelling I and age and Adding. Table III shows these correlations.

TABLE III

Relations between Test Scores and Age, Clerical Experience and Education

TEST	AGE	CLERICAL EXPERIENCE	EDUCATION
Pressey education tests	.105	.145	.416
Finger dexterity	.161	.118	−.013
Tweezer dexterity	.106	.102	.195
Spatial relations	−.054	−.046	.025
Clerical number checking	.164	.312	.282
Clerical name checking	.157	.301	.168
Feature discrimination 2A	.109	.154	.092
Feature discrimination 3A	.094	.117	.256
Feature discrimination 4A	.109	.043	.123
Feature discrimination 5A	.010	.081	.264
Number cancellation	.158	.050	.336
Letter cancellation	.148	.108	.229
Letter substitution	.135	.154	−.049
Number substitution	.101	.126	.205
Verifying arithmetic	.083	.141	.156
Spelling I	.237	.251	.368
Spelling II	.164	.258	.388
Number group checking	.193	.197	.149
Adding	.221	.196	.095

Probable error .054—correlation must be .22 to be significant.

The only correlations indicating even a slight relationship between clerical experience and test scores are those for the two clerical checking tests and the two spelling tests. Such relationships do not necessarily mean that those with more clerical experience are in a better position to score high in these tests because of that experience; rather, they mean that those who are successful clerical workers, as evidenced by many years of employment, have better clerical and spelling abilities. The relationships, however, are so slight that little weight can be placed in them.

As might be expected, academic ability is more closely related to years of schooling than to any of the tests. Those with higher academic ability ratings tend to stay in school and complete more grades than those with lower ratings. Two of the feature discrimination tests correlate better than chance with years of schooling. Since these two correlations are not for two consecutive forms, the relationship does not indicate that the more difficult tests are more affected by schooling.

The Arithmetic Test is not related to education since all individuals with an eighth grade education can answer all of the problems. Since this is a group of clerical workers who have had experience with numerical calculations it would be expected that schooling in and of itself would exert little effect.

The two spelling tests are related to schooling; those with more schooling score somewhat higher on the spelling tests. Since schooling is related to academic ability, it is likely that those who spell better are those who not only attended school longer but also have better academic ability and retained better what they learned in these grades. All other relationships in this table are low and possibly are due to chance.

When the relationships between age, clerical experience, and education are considered, the correlations show that clerical experience is not related to years of schooling, $-.015$, but there is a marked relationship between age and clerical experience as indicated by a coefficient of $+.791$. Those who are older have had more clerical experience than those who are younger.

This signifies that the older members of the group are persons who have remained in clerical work and have accepted it as a permanent vocation. There is a negligible relationship between age and years of schooling, the correlation being +.155.

INTERCORRELATIONS AND INTERPRETATIONS OF THE RELATIONSHIPS AMONG THE TESTED ABILITIES

One of the most important considerations in any study of the analysis of mental abilities is the inquiry into the intercorrelations among these abilities. All degrees of relationship are important, from the chance relations indicating the uniqueness of the variable in respect to that ability to the correlations which indicate some or a high degree of relationship.

Table IV gives the coefficients of correlations among the nineteen variables for this group of women clerical workers. The correlations to the right of the diagonal are the uncorrected correlations, and those to the left of the diagonal are the correlations corrected for attenuation. These corrected coefficients can serve only as indications of the possible correction, since the reliabilities upon which the corrections are based, are not all calculated for the same group, a group of women clerical workers.

There are ten correlations in Table IV which are greater than +.50. These most important relationships may be summarized as follows:

Number and letter cancellation	.84	corrected	.98
Spelling I and II	.70	"	.73
Clerical number and name checking	.66	"	.69
Clerical names and spelling II	.65	"	.68
Clerical names and spelling I	.63	"	.66
Number group checking and adding	.60	"	.63
Clerical number checking and spelling I	.53	"	.57
Feature discrimination 4A and 5A	.53	"	.56
Letter and number substitution	.52	"	.52
Clerical number checking and arithmetic	.51	"	.54

The highest correlations indicate what "single factorists" refer to as group factors.

TABLE IV

*Corrected and Uncorrected Intercorrelations for the Nineteen Ability Tests**

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1.		.091	.119	.227	.253	.411	.325	.271	.283	.375	.116	.142	.328	.221	.332	.453	.450	.126	.202
2.	.104		.331	.245	.255	.185	.051	.242	.021	.151	.158	.086	.162	.222	.167	.092	.027	.228	.173
3.	.136	.390		.269	.309	.214	.057	.297	.109	.202	.338	.185	.160	.177	.134	.192	.076	.160	.220
4.	.254	.282	.309		.091	.136	.272	.296	.320	.435	.276	.259	.168	.217	.190	.227	.119	.140	.262
5.	.277	.288	.349	.100		.655	.187	.376	.210	.382	.359	.416	.369	.462	.507	.528	.459	.479	.436
6.	.441	.204	.237	.146	.693		.288	.375	.288	.421	.400	.370	.431	.465	.430	.629	.653	.488	.452
7.	.347	.056	.062	.292	.197	.297		.473	.344	.247	.231	.208	.249	.206	.214	.369	.371	.282	.230
8.	.290	.256	.328	.319	.397	.387	.487		.355	.433	.347	.444	.165	.346	.247	.330	.334	.355	.350
9.	.303	.023	.126	.372	.222	.298	.354	.367		.529	.287	.193	.241	.230	.178	.255	.155	.138	.282
10.	.408	.169	.226	.475	.428	.442	.258	.455	.556		.312	.304	.240	.306	.241	.458	.328	.230	.303
11.	.135	.189	.405	.323	.412	.449	.258	.388	.322	.355		.837	.430	.366	.270	.403	.318	.482	.493
12.	.158	.098	.207	.291	.458	.398	.223	.478	.208	.332	.978		.301	.434	.369	.386	.383	.425	.482
13.	.347	.177	.174	.179	.387	.443	.253	.169	.247	.249	.477	.320		.517	.380	.349	.469	.333	.370
14.	.223	.241	.193	.230	.482	.474	.210	.352	.234	.317	.405	.461	.521		.342	.348	.349	.389	.411
15.	.357	.185	.148	.206	.538	.463	.217	.256	.184	.253	.304	.399	.388	.350		.407	.416	.302	.367
16.	.491	.114	.214	.254	.565	.659	.385	.346	.267	.487	.458	.421	.360	.359	.427		.696	.391	.414
17.	.485	.030	.083	.129	.488	.680	.385	.347	.161	.346	.358	.415	.482	.358	.434	.733		.359	.368
18.	.137	.257	.181	.154	.518	.516	.297	.375	.145	.247	.554	.468	.348	.406	.320	.418	.382		.600
19.	.214	.189	.240	.262	.456	.462	.233	.357	.288	.314	.546	.513	.373	.415	.376	.429	.378	.628	

* Uncorrected correlations are to the right of the diagonal and corrected to the left of the diagonal. Corrected correlations were found by applying the correction for attenuation and using Spearman-Brown reliabilities with the exception of the Pressey Test for which the correlation for comparable forms without correction was used.

KEY TO TABLE

1. Average of the two Pressey Tests
2. Finger dexterity
3. Tweezer dexterity
4. Spatial relations
5. Clerical number checking
6. Clerical name checking
7. Feature discrimination 2A
8. Feature discrimination 3A
9. Feature discrimination 4A
10. Feature discrimination 5A
11. Number cancellation
12. Letter cancellation
13. Letter substitution
14. Number substitution
15. Verifying arithmetic
16. Spelling I
17. Spelling II
18. Number group checking
19. Adding

Academic ability, as measured by the Pressey Tests, is most closely related to the verbal abilities, or to such tests as clerical name checking and spelling. Academic ability is related to clerical abilities and to the feature discrimination tests, but it is relatively independent of the simple dexterities, cancellation, and number group checking.

Finger dexterity is more closely related to tweezer dexterity than to any other test, but this correlation is only +.33. Finger dexterity is more related to tweezer dexterity, spatial and numerical abilities, than to academic or verbal abilities of which it is relatively independent.

Tweezer dexterity is unrelated to academic ability, number and letter substitution, arithmetic, spelling, and number group checking tests. Number cancellation, finger dexterity, and clerical number checking are associated with tweezer dexterity. It is apparent that tweezer dexterity is more closely related than finger dexterity to the clerical tests.

The Minnesota Spatial Relations Test, a measure of mechanical ability, is more closely related to the most complex feature discrimination test than to any other test included in this battery. It is also somewhat related to finger and tweezer dexterity which indicated that the three performance tests are associated. Spatial ability is definitely independent of clerical ability. Table V gives the relationships between the spatial relations test and the other tests in the battery with academic ability held constant.

Clerical number checking, although most highly correlated with clerical name checking, shows definite relations to the other clerical tests. Clerical number checking is more highly related to finger and tweezer dexterity, letter cancellation, and arithmetic than clerical names is correlated with these tests. Partial correlations with academic ability constant indicate that the only test which is more closely connected to clerical number than to name checking is the verifying arithmetic test. The only test which is definitely unrelated to clerical number checking is spatial relations.

TABLE V

Partial Correlations between Other Variables with Academic Ability Constant

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1.		.331	.051	.242	.021	.151	.158	.086	.162	.222	.167	.092	.027	.228	.173
2.	.322		.057	.297	.109	.202	.338	.185	.160	.177	.134	.192	.076	.160	.220
3.	.022	.019		.473	.344	.247	.231	.208	.249	.206	.214	.369	.371	.282	.230
4.	.226	.277	.423		.355	.433	.347	.444	.165	.346	.247	.330	.334	.355	.350
5.	−.005	.078	.278	.301		.529	.287	.193	.241	.230	.178	.255	.155	.138	.282
6.	.122	.171	.142	.369	.476		.312	.304	.240	.306	.241	.458	.328	.230	.303
7.	.148	.328	.205	.330	.266	.290		.837	.430	.366	.270	.403	.318	.482	.493
8.	.074	.174	.173	.427	.161	.272	.836		.301	.434	.369	.386	.383	.425	.482
9.	.140	.129	.159	.083	.163	.133	.418	.272		.517	.380	.349	.469	.333	.370
10.	.213	.156	.145	.304	.178	.245	.351	.417	.482		.342	.348	.349	.389	.411
11.	.146	.100	.118	.172	.093	.132	.246	.344	.304	.292		.407	.416	.302	.367
12.	.057	.171	.263	.240	.148	.347	.394	.364	.237	.284	.304		.696	.391	.414
13.	−.016	.025	.267	.246	.033	.191	.299	.361	.381	.287	.317	.617		.359	.368
14.	.219	.151	.257	.336	.107	.198	.474	.414	.311	.372	.278	.376	.340		.600
15.	.158	.201	.177	.312	.238	.458	.484	.468	.328	.383	.324	.368	.352	.592	

Correlations to the right of the diagonal are uncorrected intercorrelations; those to the left of the diagonal are the partial correlations, with academic ability constant.

1. Finger dexterity
2. Tweezer dexterity
3. Feature discrimination 2A
4. Feature discrimination 3A
5. Feature discrimination 4A
6. Feature discrimination 5A
7. Number cancellation
8. Letter cancellation
9. Letter substitution
10. Number substitution
11. Verifying arithmetic
12. Spelling I
13. Spelling II
14. Number group checking
15. Adding

Clerical number checking and the two spelling tests are definitely related to clerical name checking. With academic ability held constant, clerical name checking is more closely correlated with feature discrimination 2A and 4A and the two spelling tests than is clerical number checking. The other clerical tests show some degree of relationship to name checking, although academic ability is somewhat more closely related to name than to number checking. Clerical name checking is relatively independent of finger dexterity and spatial relations. Even after academic ability is held constant, there is a relationship of .58 between clerical name checking and spelling II.

Feature discrimination 2A, the simplest among the discrimination tests, is more closely related to feature discrimination 3A than to any other test. It is also correlated with academic ability and the spelling tests, which correlations show a considerable decrease when academic ability is held constant. Finger and tweezer dexterity, letter substitution, and clerical number checking are definitely unrelated to this test of observation and comparison.

Feature discrimination 3A like the other tests of observation and comparison is more closely related to other feature discrimination tests, in this case 2A and 5A, than to any other tests. All of the clerical series are correlated positively with feature discrimination 3A with the exception of letter substitution which is relatively independent of this test.

Feature discrimination 4A is most closely associated with other discrimination tests and spatial relations. The only test with which 4A is more closely connected than 5A is feature discrimination 2A. Finger and tweezer dexterity, clerical number checking, letter cancellation, arithmetic, spelling II, and number group checking show little indication of being related to the 4A discrimination test.

Feature discrimination 5A is more closely related than feature discrimination 4A to the clerical and spatial abilities. This test, the most complex in the discrimination series, is more closely related to spatial and to the other feature discrimina-

tion tests than is clerical name checking. It is also more closely associated than spatial relations to most of the clerical tests and to academic ability. Finger and tweezer dexterity are the only tests in this battery which are relatively independent of this test in feature discrimination.

Feature discrimination tests 3A and 5A are closely related to the clerical tests; 4A and 5A are closely associated with spatial ability, and 2A and 5A with academic ability. These tests, designed to measure the ability to observe and compare differences, measure not only academic ability, as stated by Greene, but abilities which are important in the clerical and spatial tests. Although all four discrimination tests are similar excepting for their complexity, the relationship between any two of them is not as high as those between two spelling, clerical checking, or cancellation tests. Each form measures an ability somewhat different from the preceding.

The two cancellation tests are very closely related as shown by the correlation of +.84, or corrected +.98. The clerical tests particularly number group checking, adding, substitution, and clerical name and number checking tests are related to the cancellation tests. Two tests which are definitely unrelated to cancellation are finger dexterity and academic ability. The type of motor ability required in tweezer dexterity is more closely related than finger dexterity to speed in cancellation.

The two substitution tests are more closely related to each other than to any other test, as is evidenced by the intercorrelation of +.52. Both tests are associated significantly with clerical and academic ability. Letter substitution is relatively independent of finger and tweezer dexterity, spatial relations, and 3A of the discrimination tests. Number substitution is unrelated to tweezer dexterity.

Arithmetical ability, as measured by the Verifying Arithmetic test, is more closely dependent upon clerical number checking than upon any other test. This relationship is little affected by the presence of academic ability in the correlation while other relationships between arithmetic and clerical name

checking or the spelling tests are greatly decreased when academic ability is held constant. The only tests which are definitely unrelated to arithmetic are finger and tweezer dexterity, spatial relations, and feature discrimination 4A.

The two spelling tests are more closely associated with each other than with any other test. They are, however, very closely related to clerical name checking. The higher correlations between spelling and clerical name checking cannot be explained by a common factor of speed since all of the tests in the battery are time tests. To test the importance of speed in such a relationship, 77 vocational high school students were given the clerical test and a 50 word spelling test. Speed was not a factor in the spelling test, since all words were pronounced singly and enough time was allowed for each student to write each word. The correlation between this test in commercial spelling and clerical number checking is +.22, while the relationship between spelling and clerical name checking is +.55.

Clerical name checking is more highly correlated with other clerical tests than are the spelling tests. The spelling tests are significantly related to clerical and academic abilities. Spelling I is independent of finger dexterity and feature discrimination 2A. Finger and tweezer dexterity, spatial relations, and 4A are unrelated to the more complex spelling test.

Number group checking and adding are rather highly correlated, as shown by the correlation of +.60 or of +.59 with academic ability held constant. Number group checking is related to all the clerical tests and is independent of academic ability, tweezer dexterity, spatial relations, and feature discrimination 4A. Number group checking, then, is measuring a clerical factor which is free from the influence of academic, dexterity, and spatial factors.

Although number group checking and adding are associated, the adding test shows more relationship to academic ability, tweezer dexterity, spatial relations, the more complex discrimination tests, and arithmetic tests than does number group

checking. Adding is related to the clerical tests, academic, and spatial abilities, but is independent of finger dexterity.

From the discussion of the intercorrelations, it is evident that there are several factors operating, among which seem to be academic ability, clerical aptitude, spatial ability, and possibly a dexterity factor. There is the additional probability that clerical ability may include a numerical and a verbal factor. Factor loadings and tetrad analysis will be required to provide an analysis of these abilities.

The above conclusions have been drawn from an analysis of each test when considered alone. Table VI shows some of the relationships which are present when the tests are grouped according to the following plan:

Dexterity: Finger and Tweezer.

Feature discrimination: 2A, 3A, 4A, 5A.

Verbal ability: Clerical: Clerical name checking, letter cancellation, letter substitution, and spelling II.

Numerical ability: Clerical: Clerical number checking, number cancellation, number substitution, arithmetic, number group checking, and addition.

The average intercorrelation for all of the variables of +.27 indicates the tendency for all desirable traits to be positively correlated. Clerical experience and education are independent of all the test groups. Academic ability is related to the feature discrimination and verbal (clerical) tests, but is relatively independent of dexterity and numerical (clerical) tests. The feature discrimination tests are most closely related to spatial ability, then academic, clerical-verbal, and numerical abilities. Dexterity is not related to academic, verbal-clerical, numerical-clerical or feature discrimination tests. Dexterity is slightly related to spatial ability. Verbal-clerical and numerical-clerical abilities are independent of spatial ability, but spatial ability is somewhat related to dexterity, academic, and feature discrimination abilities.

There is little evidence for the isolation of a verbal and numerical factor from the intercorrelations for these test

TABLE VI

Average Correlations between Various Groups of Tests

	r
Clerical experience and feature discrimination	.099
Clerical experience and dexterity	.110
Clerical experience and verbal (clerical)	.156
Clerical experience and numerical (clerical)	.138
Education and feature discrimination	.184
Education and dexterity	.091
Education and verbal (clerical)	.191
Education and numerical (clerical)	.155
Academic ability and feature discrimination	.314
Academic ability and dexterity	.105
Academic ability and verbal (clerical)	.333
Academic ability and numerical (clerical)	.208
Feature discrimination and dexterity	.141
Feature discrimination and verbal (clerical)	.288
Feature discrimination and numerical (clerical)	.269
Feature discrimination and spatial	.331
Dexterity and spatial	.257
Dexterity and verbal (clerical)	.137
Dexterity and numerical (clerical)	.212
Spatial and verbal (clerical)	.171
Spatial and numerical (clerical)	.196
Clerical verbal and numerical	.438
Clerical verbal and verbal	.435
Clerical numerical and numerical	.418

groups. The verbal tests correlated with the verbal tests no more highly than did the verbal tests with the numerical group. The numerical tests were also no more closely related to the numerical tests than they were to the verbal tests. These intercorrelations in the +40's indicate a fair degree of relationship between the various clerical tests, whether they be of a verbal or of a numerical content.

Although the tests when grouped show no evidence for a verbal and numerical factor, two of the tests, those of arithmetic and spelling ability, which unlike the other tests, truly represent numerical and verbal content, do give some indications of the differentiation of numerical from verbal facility.

SUMMARY AND CONCLUSIONS

The purpose of this study was to make a comprehensive analysis of the Minnesota Vocational Test for Clerical Workers. Earlier results have indicated that the Minnesota Test is a reliable and valid measure of clerical aptitude, and is relatively free from the influence of clerical training and experience. Both the clerical number and name checking tests are relatively independent of academic ability when homogeneous groups are considered. For more heterogeneous groups, only clerical number checking is independent of academic ability. The clerical test differentiates clerical workers from workers in the general population and from workers in various non-clerical occupations.

The four problems which have been singled out for study are:

(1) to discover what aptitudes the clerical test is measuring and what traits are closely related to it.

(2) to determine whether the clerical test is measuring a unique trait which is relatively independent of other tested abilities.

(3) to determine the value of multiple factor and tetrad analysis techniques in the study of the overlapping among abilities and of the number of factors involved in a test battery which measures academic, clerical, dexterity, and mechanical abilities.*

(4) to evaluate the contributions and the limitations of factor analysis techniques, in a vocational testing program.*

An attempt has been made in this study to meet the usual criticisms made against studies in mental analysis and the interrelationships among abilities. A group of young women clerical workers, homogeneous as to sex, age, and general occupation has served as the experimental group. The tests were used only after they had shown themselves to be reliable measuring instruments and capable of giving a normal distribution of scores.

* The last two problems will be discussed in a subsequent article in this journal.

To study the first problem, namely, that of the nature of the test and the abilities to which it is related, tests were selected which supposedly measured relatively simple clerical abilities and were revised to meet the requirements of reliability. These tests include number and letter cancellation, letter and number substitution, arithmetic, spelling, number group checking, and addition. Four feature discrimination tests from the Michigan Non-Verbal Series devised by E. B. Greene were also used because they supposedly measured an ability to observe and compare an ability which might be closely related to the clerical test.

The uniqueness of the clerical test was studied by observing its relation to the Pressey Education Tests, the O'Connor Finger and Tweezer Dexterity Tests, and the Minnesota Spatial Relations Test, which tests were used by the Minnesota Employment Stabilization Research Institute in its testing program.

RESULTS

All of the tests with the exception of number cancellation, have reliabilities of at least +.85 which are sufficiently high for purposes of individual diagnosis. The number cancellation test is reliable enough to determine group differences.

The average intercorrelation for all of the variables is +.27 which indicates the tendency for desirable traits to be positively correlated. The uncorrected correlations range from .00 to +.84; the highest correlation is between number and letter cancellation.

Correlational analysis seems to indicate that this test battery is measuring at least four relatively independent factors, those of academic, clerical, spatial, and dexterity abilities. There is also the possibility that the clerical number and name checking tests are measuring two factors, one a numerical ability, and the other a verbal ability. There are also indications that the two dexterity tests, though related, are measuring relatively specific abilities. The feature discrimination tests seem

to be measuring abilities which are important in academic, clerical, and spatial relations tests. Although spatial relations is more closely related to finger and tweezer dexterities than are the feature discrimination tests, the feature tests are more closely associated than spatial relations to clerical and academic abilities. When the feature discrimination tests are considered in one group the verbal-clerical tests separated from the numerical-clerical tests and the two dexterity tests combined, the relationships give a composite picture of the many correlations.

The correlations indicate that both clerical experience and education are independent of the abilities measured by these various test groups. Academic ability is significantly related to the verbal-clerical and feature discrimination tests but is independent of the numerical-clerical tests and dexterities. The feature discrimination tests are related to the spatial and clerical tests, but are independent of the dexterity tests. The dexterity tests are somewhat related to spatial ability but are independent of other test groups. Spatial ability is definitely unrelated to the clerical test groups, those dealing with letters and those dealing with numbers, no differences are present between the correlations of the verbal with the verbal and the verbal with the numerical tests. Although the arithmetic, spelling, and clerical number and name checking indicate some differences between the verbal and numerical abilities, other tests of substitution, cancellation, number group checking, and addition do not indicate such differences.

BIBLIOGRAPHY

1. Ackerson, L. A Correlation Analysis of Proficiency in Typing, Archives of Psychol., 82, 1926.
2. Andrew, D. M., and Paterson, D. G. Measured Characteristics of Clerical Workers, Bull. of the Empl. Stab. Res. Inst., Univ. of Minnesota, 1934, Vol. III, No. 1.
3. Bailor, E. M. Content and Form in Tests of General Intelligence, N. Y., Teachers College, Columbia Univ., 1924.
4. Burt, C. Distribution and Relation of Educational Abilities, 1916, London County Council, p. 46.

5. BURTT, H. E. "Employment Psychology in the Rubber Industry," J. Appl. Psychol., 1920, 4, 1–17.
6. DVORAK, B. J. Differential Occupational Ability Patterns, Bull. of Empl. Stab. Res. Inst. Univ. of Minn., 1935, Vol. III, No. 8.
7. GREEN, H., BERMAN, I. R., PATERSON, D. G., and TRABUE, M. R. A Manual of Selected Occupational Tests for Use in Public Employment Offices, Bull. of Empl. Stab. Res. Inst., Univ. of Minnesota, 1933, Vol. II, No. 3.
8. GREENE, E. B. Manual for the "Michigan Non-Verbal Series," Ann Arbor, Edwards Bros., 1931.
9. HIGHSMITH, J. A. Relation of the Rate of Response to Intelligence, Psychol. Monog., 1925, 34, No. 3, 1–33.
10. LINK, H. C. Employment Psychology, N. Y., Macmillan Co., 1919, 77–111 and 417.
11. MOORE, H. "The Institute's Clerical Test in America," Human Factor, 1933, 7, 407–409.
12. MUNRO, M. S., and RAPHAEL, W. "Selective Tests for Clerical Occupations," J. Nat'l. Inst. Industr. Psychol., 1930, 5, 127–137.
13. PATERSON, D. G. (Editor). Research Studies in Individual Diagnosis, Univ. of Minnesota, Bull. of Empl. Stab. Res. Inst., 1934, Vol. III, No. 4.
14. PIDDINGTON, R. "Psychological Tests for Clerical Workers," Australasian J. Psychol., 1930, 8, 59–68.
15. THURSTONE, L. L. "Standardized Tests for Office Clerks," J. Appl. Psychol., 1919, 3, 248–251.

PERSONALITY AND ATTITUDINAL SIMILARITY AMONG CLASSROOM FRIENDS

R. PINTNER, G. FORLANO AND H. FREEDMAN
Teachers College, Columbia University

AN objective study of social relationships constitutes an important but difficult task. Of the many forms of social relationships, that of friendship has received much attenion. Our chief purpose here is to use objective tests, in order to ascertain the degree of similarity of certain personality traits among children and their school friends.

A number of interesting psychological and sociological investigations have been made during the past twelve years. A brief summary of the outstanding studies in the field will be presented before proceeding with a description of this study.

During the preschool period, most of the investigators, such as Challman (3), Beaver (2), Salusky (9), Parten (8), Zaluzhni (13), agree that the amount of friendship increases with age. Challman studied thirty-three children, 17 boys and 16 girls, aged 27 to 59 months, for a period of six months. He recorded the number of times a child was seen to play with another child. The relationship of chronological age, mental age, intelligence quotient, height, attractiveness of personality, degrees of extroversion, cooperation or organized action and physical activity, with friendship was studied. He concluded that boys show preference for boys and girls for girls. Boys pick friends who are similar with respect to degree of cooperative activity, degree of physical activity and chronological age. Girls show the same tendency but to a lesser degree, especially in physical activity. Boys are more discriminative in their friendships than the girls. Neither the boys nor girls preferred friends who were similar to themselves in

mental age, intelligence quotient, height, attractiveness, personality or degree of extroversion.

Green (5), and Chevaleva-Janovskaja (4) agreed with Challman that preschool children tend to seek friends of their own age. Hagman's (6) findings are also similar. She did not find that preschool children tend to pick friends who are similar in mental age or intelligence quotient. Similarities in weight, height, number of days children are in school together, social stimulus index and social reaction index, do not influence frienship. However, association outside of school does have an effect.

In so far as these studies of preschool children are concerned, we do not find any marked tendency for children to choose friends who are similar to themselves in personality traits.

Coming now to studies dealing with children above the preschool age, we note that chronological age continues to play an important rôle in the formation of friendships. On this level, mental age and intelligence quotient begin to assume importance as the basis for friendship. Almack (1) studied 387 children from grade IV to VII inclusive of the public schools of San José, California. The National Intelligence Test was given. The child was asked to select a friend with whom he would like to work on a problem, and a friend with whom he would like to go to a party. The following coefficients of correlation were obtained between the child and his friends:

	C.A.	I.Q.	M.A.
Work Associate	.531	.408	.541
Play Associate	.496	.317	.304

The results show that there was a tendency for the child to choose a friend more similar to himself in picking an associate for work than for play. It also shows a substantial resemblance between friends in intelligence and chronological age.

Wellman (11) compared the similarity of friends in chronological age, intelligence quotient, achievement in scholarship, extroverted tendencies, height and physical achievement. She tested 63 boys and 50 girls in grades VII, VIII, and IX. The children were observed for a period of five months in the freer activities of the school. She found that boys and their friends are more alike in height, intelligence quotient, and chronological age, and less alike in extroversion, scholarship and mental age. Girls and their friends were found to be more alike in scholarship than in any other characteristic and less alike in height.

Williams (12) studied 84 adolescent delinquents whose chronological ages ranged from 12 to 17 years and whose mental ages ranged from 9 years to 15 years, 6 months. He found a slight tendency for the boys to choose friends who were older but of about the same mental age.

Jenkins (7) found that there is no tendency for a child to pick friends older or younger than himself. The child chooses friends of approximately the same intelligence. In a study of 180 junior high school children and their friends, he reported the following correlations:

Chronological Age	.47 ± .04
Mental Age	.423 ± .04
Intelligence Quotient	.332 ± .04

A recent study involving college students and using tests of attitudes as well as tests of intelligence was made at De Pauw University by Vreeland and Corey (10). Thirty pairs of the same sex were selected. It was found that neither intelligence nor academic grades constitute factors of importance in friendship selection among college students. The following correlations were obtained for the college pairs:

Freshman Intelligence Scores	(15 pairs)	−.302 ± .15
Thurstone Neurotic	(30 pairs)	.323 ± .12
Moss Social Intelligence	(30 pairs)	.403 ± .10
Droba, Attitude Toward War	(30 pairs)	−.016 ± .12
Watson Test of Public Opinion	(30 pairs)	.090 ± .12

Vreeland and Corey concluded that, "Only two of the tests given showed a positive correlation of any consequence, the tests for neuroticism and social intelligence. The study suggests that intelligence and scholarship as well as prejudice regarding intellectual questions, not a part of the student's immediate experience, were relatively unimportant but that degree of neuroticism and social intelligence play some rôle in the selection of friends."

Object of this Study.

The object of this study is to ascertain the amount of similarity of certain attitudes and personality traits among school children and their friends. Do birds of a feather flock together?

The Subjects.

All the tests were given to public elementary school children whose grade range was from V to VIII and whose chronological age range was 9 to 16 years. The subjects were selected from three schools in the City of New York, and one school in a small community in New York State. There were considerable differences in socio-economic background and natio-racial origin among the populations of the various schools. Forty per cent of the population in School A was colored, the remainder was composed of children of Porto Rican, Spanish, Italian, and Irish parentage. The population of School B was largely composed of children of German origin. There were no colored children in this school. Only girls attended School A. The population in School D was composed of Jewish, Italian, and Irish children.

	BOYS	GIRLS	TOTAL
School A	0	134	134
School B	85	75	160
School C	68	68	136
School D	205	184	389
Total Number of Children	358	461	819

The tests used in this study were as follows:

1. Cultural Attitudes Toward Animals.
2. Cultural Attitudes Toward Nature.
3. Cultural Attitudes Toward Art.
4. Cultural Attitudes Toward Music.
5. Maller Character Sketches.[1]
6. Pupil Portraits.[2]
7. Personality Outlines.[3]
8. Pupil Questionnaire.
9. General Opinion Test.
10. Woodworth-Matthews Questionnaire.[4]
11. Pintner Rapid Survey Intelligence Tests. Grades IV–VIII.

Description of the Tests Used.

Almost all of the tests are of the self-descriptive type in which the pupil indicates by checking or marking Same or Different what he thinks about his own feeling, attitudes, or personality characteristics. Tests 5, 6, 10, and 11 are published standardized tests. All the other tests are in process of construction. Since the published tests are available to all who wish to examine them, no items are given here. The sample items of the unpublished tests follow:

1. Cultural Attitudes—Animals.
   S. D. 1. This child thinks that most animals are bad.
   S. D. 2. This child thinks that people are doing something wrong when they kill birds.

[1] Maller, J. B. Character Sketches, Bureau of Publications, Teachers College, Columbia University, 1932.

[2] Pintner, R., Maller, J. B., Forlano G., Axelrod, H. Pupil Portraits, Journal of Ed. Res., 1935, 28, 334–346. Test blanks published by Bureau of Publications, Teachers College, Columbia University, New York City.

[3] Test in preparation by Pintner, R., Loftus, J. J., Forlano, G., Alster, B. Teachers College, Columbia University, New York City.

[4] Woodworth-Matthews Personality Questionnaire. C. H. Stoelting Co., Chicago, Illinois.

2. Cultural Attitudes—Nature.
   S. D. 1. This child thinks that collecting plants and flowers is not as interesting as playing games.
   S. D. 2. This child would walk a mile to see a beautiful sunset.
3. Cultural Attitudes—Art.
   S. D. 1. This child thinks that it is a waste of time to look at pictures.
   S. D. 2. This child thinks that all people should be taught how to paint and draw.
4. Cultural Attitudes—Music.
   S. D. 1. This child thinks that music doesn't count anyway because we are not marked for it at school.
   S. D. 2. This child would like to play music with other children.
7. Personality Outlines.

   This test was revised while the experiment was being conducted. School B used the revised form of the experimental edition used in School A. It consists of items purporting to measure ascendance-submission, extroversion-introversion, superiority-inferiority, and emotional stability. In the revised form the superiority-inferiority test does not appear. Sample items of each subtest are given.

   A. Ascendance-Submission.
      Same Different 1. I try to be the first to get on the street car.
      Same Different 2. I stick to what I've said, even if other people don't like it.
   B. Emotional Stability.
      Same Different 1. I can't forget a wrong that's been done me.
      Same Different 2. I worry about getting sick.
   C. Extroversion-Introversion.
      Same Different 1. I would like to go to school early because I have friends waiting for me.
      Same Different 2. I would like to be in assembly plays.

8. Pupil Questionnaires.

This test consists of short impersonal statements very much like those of the Woodworth-Matthews Questionnaire.

Yes No 1. Are you often frightened in the middle of the night?

Yes No 2. Do you ever feel like running away from home?

9. General Opinion Test.

At School B, the General Opinion Test was substituted for the Pupil Questionnaire. An experimental form of the General Opinion Test was used at School C. The results at School C did not justify our using this test again until it had undergone a thorough revision. The revised form was given at School B.

Another factor employed was a measure of popularity. This was determined by the child's selection of his school companions as his friends. Two types of popularity scores were employed. The unweighted score (A) represents the total number of times the child was chosen by his friends regardless of placement. The weighted score (B) depended on the order in which a child was chosen by his friends. The weights were 3, 2, 1, according as he was ranked as "best friend," second or third.

Reliability of the Tests.

The reliability of these tests was determined by the correlation between the odd and even items. The Spearman Brown prophecy formula was then employed to obtain the reliability coefficient for the total test. The number of cases used was 100. The coefficients of reliability follow on page 55.

For the first six tests, for the emotional stability items on the Personality Outlines, and for the total scores on the Personality Outlines, the reliabilities are quite satisfactory. For the other parts of the Personality Outlines, the reliabilities are very low.

	r HALF	r WHOLE	ITEMS ON TEST
1. Cultural Attitudes Toward Animals	.73	.84	50
2. Cultural Attitudes Toward Nature	.80	.89	50
3. Cultural Attitudes Toward Art	.72	.84	50
4. Cultural Attitudes Toward Music	.81	.90	50
5. Character Sketches	.69	.84	100
6. Pupil Portraits—Total	.80	.88	100
School Adjustment	.68	.81	73
Home Adjustment	.47	.64	27
7. Personality Outlines—Total	.71	.83	120
Emotional Stability	.65	.79	30
Superiority	.43	.60	30
Extroversion	.30	.46	30
Ascendance	.30	.14	25

Procedure and Administration of the Tests.

In Schools A, B, and C, every child was given an index card on which he listed his three best friends in the class, in the order of his preference. It was necessary to so limit the child's choice of friends to his classroom in order that we might secure the friends' scores as well as those of the child. There was no such limitation made in the selection of friends, in School D, as every class from grades 4A to 8B was tested.

In giving the personality and attitude tests certain explanations were made to the children. Before beginning work on these tests the children were told that these tests had no "right answers." The children were further told that neither teacher nor principal would see the papers. The examiner explained the meaning of any words or phrases which the child did not understand. The child raised his hand during the test whenever he failed to understand a statement. He was told that it was not a test in which he would receive a mark. The teacher was not in the room whenever a test was in progress.

The testing procedure was similar in Schools A, B, C, and D. Each test took about thirty minutes to administer; some children finished in twenty minutes, a few took forty minutes.

A preliminary survey with experimental editions of the tests, was made at School C. School B was used as a check on Schools A and C. Revised forms of the Cultural Attitudes Tests, Personality Outlines Test, and General Opinion Test, were used at School B. School D, at which the revised forms of the Cultural Attitudes and Personality Outlines were used, was a final check on the results for Schools A, B, and C. The results are not given for boys and girls separately, since it was found that the children chose friends of the same sex although there were no restrictions as to choice on this point.

A summary card was prepared for each child. The child's three friends were listed on the card. The friends' scores appeared on the reverse side of the card. The correlation technique seemed to be the most appropriate for our problem, namely, to find out whether classroom friends are similar with respect to personality and behavior characteristics. The following pertinent correlations were obtained:

1. The correlation between the child's score and the scores of his three friends for each test.
2. The correlation between the child's chronological age and the ages of his three friends. (Schools B and D.)
3. The correlation between the child's mental age and the mental ages of his three friends. (Schools B and D.)
4. The correlation between the child's intelligence quotient and the quotients of his three friends. (Schools B and D.)
5. The correlation between the child's popularity score and the scores of his three friends.

To supplement the correlations outlined above correlations were obtained between the child and his best friend, and the child and his third friend, on each of the tests or factors concerned. It was deemed advisable to ascertain the effect of limiting the range of the population on the results. Therefore in Schools A and B correlations between (1) the child and his three friends, (2) the child and his best friend, (3) the child and his last choice, were obtained for the sixth-year class.

In calculating the correlations between the child and his three friends, three scores were entered for each child. The X axis represented the child's score, the Y axis represented the friends' scores. The Pearson product moment coefficient of correlation was employed.

The Attitude and Personality Scores.

Results in School C.

Experimental editions of the tests were given in School C. Four tests were given to 5A, 5B, 6A, and 6B. Three correlations were computed for each test. The number of children tested was 136.

TABLE 1

Correlation of Scores Between Child and His Three Friends at School C

TESTS	CHILD AND THREE FRIENDS	CHILD AND BEST FRIEND	CHILD AND LAST FRIEND
General Opinion Test	.35	.34	.27
Character Sketches	.28	.39	.38
Woodworth-Matthews	.10	.10	.16
Pupil Portraits	.16	.19	.10
Popularity—A	.09	−.03	.15
Popularity—B	.12	−.04	.07

In Table 1, the correlations with the exception of two are positive. Though the correlations fluctuate, there seems to be a slight tendency for the child to resemble his friends. It is interesting to note the general tendency for the correlation between the child and his best friend to be higher than the correlation between the child and his last friend, but this does not hold for the popularity scores.

Results in School A.

Table 2 shows the correlations between the child and his three friends, the child and his best friend, and the child and his third friend for grades 5B to 8A; and the correlations between the child and his three friends, his best friend and his third choice for 6B in School A. The number of children tested was 134.

TABLE 2

Correlations of Scores Between Child and His Three Friends at School A

TESTS	5B, 6B, 7B AND 8A			6B ONLY		
	Child and Three Friends	Child and Best Friend	Child and Last Friend	Child and Three Friends	Child and Best Friend	Child and Last Friend
Cultural Att.—Animals	.16	.16	.06	.25	.08	.45
Cultural Att.—Nature	.07	.14	.08	−.16	.17	−.25
Cultural Att.—Art	−.10	.03	.01	−.05	.07	−.37
Cultural Att.—Music	−.10	−.14	−.07	−.04	.03	.08
Character Sketches	.04	.19	.03	.08	.05	.03
Pupil Portraits—Total	.11	.14	−.04	.19	.21	.02
School Adjustment	.09	.07	−.06	.06	.20	−.02
Home Adjustment	.16	.20	−.03	.18	.28	.01
Personality Outlines						
Ascendance	−.05	.20	.02	.10	.25	.04
Extroversion	−.05	.12	−.09	−.16	−.12	.18
Superiority	.14	.23	.03	.10	.13	−.18
Emotional Stability	.15	.29	.10	.06	.20	−.11
Pupil Questionnaire	.05	.10	−.00	−.10	.18	−.12
Popularity—A	−.01	−.05	−.01	−.09	−.31	.13
Popularity—B	−.03	−.01	−.02	−.06	−.16	.22

In Table 2, the correlations are low. They vary about zero. These correlations suggest that children do not choose friends who are very much like themselves or entirely different from themselves. The results may have been affected by limiting the child's range of friendship. He had to choose children of his own class for his friends. However, the tests in this school as well as in Schools B, C, and D were given at the end of the term so that the child knew these friends for at least five months. If we examine the correlations in class 6B, we find somewhat similar results. The cultural attitudes correlations for 6B seem to be higher in positive and negative directions than those for the entire group. In Table 2, we find higher correlations existing between the child and his best friend, than those between the child and his third friend. This result verifies that obtained in School C. In Schools A and C, the correlations of the popularity scores between the three friends, best friend and third friend seem to show that the popular child does not tend to pick the popular child for his friend. The correlations obtained for the most part are very small both positively and negatively. The correlation of the popularity score for the child and his best friend in class 6B is considerably lowered when the weighted popularity score is used.

Results in School B.

School B, unlike Schools A and D, is not situated in a populated section of the city. It is located in a suburb. The children seem to be more closely knit together than those in a crowded city school. Revised forms of the Personality Outlines and General Opinion Test were used. Table 3 shows the correlations between the child and his three friends, the child and his best friend, the child and his last friend for grades 5B–8A; and the correlations between the child and his three friends, his best friend, and third friend for Class 6A alone.

As has been noted previously the correlations in general are low. The highest correlations in the table are for chronolog-

TABLE 3

Correlations of Scores Between Child and His Three Friends at School B

TESTS	5A, 6A, 7A AND 8A			6A		
	Child and Three Friends	Child and Best Friend	Child and Last Friend	Child and Three Friends	Child and Best Friend	Child and Last Friend
Cultural Att.—Animals	.15	.10	.12	.08	.05	.01
Cultural Att.—Nature	.20	.09	.29	−.06	−.29	.19
Cultural Att.—Art	.14	.06	.24	.01	−.04	.29
Cultural Att.—Music	.07	.02	.15	−.11	−.19	−.01
Character Sketches	.02	.08	.08	−.12	−.03	−.13
Pupil Portraits—Total	.22	.36	.25	.10	.34	−.10
School Adjustment	.21	.26	.24	−.04	−.03	−.01
Home Adjustment	.16	.20	.14	.18	.12	.08
Personality Outlines—Total	.71	.75	.77	−.02	.10	.00
Ascendance	.09	.13	.10	.02	.12	.00
Extroversion	.58	.57	.57	.10	.18	−.05
Emotional Stability	.72	.74	.74	−.02	.23	.11
General Opinion Test	.20	.24	.20	.14	.24	.19
Mental Age	.62	.60	.57	.16	.17	.13
Chronological Age	.76	.78	.77	.20	.21	.09
Intelligence Quotient	.15	.10	.09	.13	.03	.17
Popularity—A	−.07	−.04	−.02	−.29	−.07	−.11
Popularity—B	−.21	−.12	−.04	−.19	−.23	−.13

ical age and mental age for the total group. These correlations drop markedly when the range is restricted to one grade, namely, grade V1. This shows the spurious nature of these correlations when a wide range of grades is employed. Chronological age remains fairly high even within the single grade range. And, we are tempted to conclude that chronological age is probably the chief determining factor in the choice of friends among children in the same school.

Results in School D.

That the restriction heretofore placed on the choice of friends might influence the correlations thus far obtained, is admitted. At School D, the choice of friends was widened. The children did not have to choose friends from their own class, as all classes from the fifth year to eighth year were tested. The results are shown in Table 4. The number of children tested in this school was 389.

In Table 4, the correlations are almost all positive. For the 5A–8B group the correlations are .10 or higher. The revised form of the cultural attitudes tests give higher correlations than the forms used in Schools A, B, and C. In general the correlation between the child and his best friend, for the total group, and for the sixth year are slightly higher than the correlations obtained for the last choice. But in general the correlations are too low to indicate definitely that a child chooses a friend similar to himself with respect to cultural attitudes and personality traits. Our former observations, that the popular child does not tend to pick the popular child for his friend, is again confirmed. The correlations again are too low to prove that the popular child chooses a popular child for a friend. We note, too, that the correlations for chronological and mental age are quite high, .76 and .62 respectively. These results substantiate those found in School B, for chronological and mental age. Another correlation of note is that of the intelligence quotient of the child and his three friends. In School B the correlations obtained between the intelligence

TABLE 4

Correlations of Scores Between Child and His Three Friends at School D

TESTS	5A TO 8B			6A TO 6B		
	Child and Three Friends	Child and Best Friend	Child and Last Friend	Child and Three Friends	Child and Best Friend	Child and Last Friend
Cultural Att.—Art	.17	.26	.13	.26	.42	.20
Cultural Att.—Music	.17	.27	.16	.32	.42	.27
Pupil Portraits—Total	.14	.26	.12	.05	.20	.12
School Adjustment	.18	.30	.18	.13	.23	.25
Home Adjustment	.02	.12	.00	.08	.21	.03
Personality Outlines—Total	.16	.19	.21	.01	.03	.02
Ascendance	.04	.14	−.03	−.04	.08	.01
Extroversion	.12	.19	.15	.02	.04	.06
Emotional Stability	.14	.18	.22	−.06	.02	.02
Mental Age	.46	.46	.45	.06	.05	.08
Chronological Age	.65	.63	.67	.20	.21	.28
Intelligence Quotient	.30	.22	.34	.11	.11	.20
Popularity—A	.13	.14	.10	.04	.00	−.08
Popularity—B	.02	−.01	.05	−.01	−.01	−.08

TABLE 5

Correlations of Scores on Attitude and Personality Tests with Popularity Scores A and B in Schools A, B, C, and D

TESTS	SCHOOL A		SCHOOL B		SCHOOL C		SCHOOL D	
	Pop. A	Pop. B	Pop. A	Pop. B	Pop. A	Pop. B	Pop. A	Pop. B
Cultural Attitudes—Animals	.17	.18	−.15	−.14				
Cultural Attitudes—Nature	−.12	−.02	−.08	−.11				
Cultural Attitudes—Art	.04	−.03	.04	.06			.02	.00
Cultural Attitudes—Music	−.15	−.09	.10	.09			−.19	−.03
Character Sketches	.09	.16	.07	.07	.31	.35		
Pupil Portraits—Total	.05	.06	.13	.14	.28	.21	.10	.24
School Adjustment	.08	.05	.10	.10			.09	.32
Home Adjustment	.06	.03	.08	.10			.08	.27
Personality Outlines—Total			.03	.06			.17	.18
Ascendance	−.12	−.05	.04	.07			.02	.00
Extroversion	.13	.10	.08	.10			.15	.17
Superiority	.06	.07						
Emotional Stability	.11	.10	.03	.09			.12	.13
General Opinion Test			.00	.03	.24	.21		
Woodworth-Matthews					.27	.28		
Mental Age			−.13	−.07			.18	.48
Chronological Age			−.06	−.01			.13	.28
Intelligence Quotient			−.09	−.06			.11	.13

quotient of the child and his three friends were low. In School D these correlations, between the IQ of the child and his friends, though low in comparison with the mental age and chronological age correlations, are slightly higher than those found in School B.

The Popularity Scores.

What personality traits go with popularity in the classroom? To answer this question the correlations between the popularity scores and the various tests were obtained. (Table 5.)

With the exception of the correlation between mental age and popularity B in School D, all the correlations are so low that for all practical purposes we may say that the scores on the above tests are not influenced by the child's popularity as measured here. A child's popularity with other children, as measured by their choice of him as a friend, does not seem to depend on those attitudes and personality traits which we have attempted to measure by means of our tests. High scores on our tests represent in general conventionally desirable traits. To possess large amounts of such desirable traits does not insure popularity in the eyes of other children. At the same time, it must be equally emphasized that large amounts of undesirable traits are not related to popularity.

SUMMARY

Several personality and attitude tests were given to 819 children in grades V to VIII in four different schools. Each child indicated his three best friends, ranking them in order. Correlations between the scores of the child and his friends are all low for all of the attitude and personality tests given. On the other hand, the correlations for age and mental ability are somewhat higher, and they are of about the same order as reported by other workers.

We must conclude, therefore, that insofar as our tests give true measures of personality characteristics, there is no ten-

dency for a child to be influenced by such characteristics in the choice of friends. It is not a negative correlation. The child does not choose a friend of opposite characteristics. It is a zero (or near zero) correlation. The friend is just as likely to differ from, as to resemble, the child in question. The positive correlations with C. A. and M. A. lead us to conclude that physical maturity and to some extent mental maturity are far more potent in influencing friendship than are the personality traits we have assumed we are measuring.

BIBLIOGRAPHY

1. ALMACK, J. C. The Influence of Intelligence on Selection of Associates. School and Society, 1922, 16: 529–530.
2. BEAVER, A. P. The Initiation of Social Contacts by Preschool Children. N. Y., Columbia University. 1932.
3. CHALLMAN, R. C. Preschool Friendships. Child Study, 1930, 8: 106–107.
4. CHEVALEVA-JANOVSKAJA, E. Les groupements spontanes d'enfants à l'age préscolaire. Arch. de psychol., 1927, 20: 219–233.
5. GREEN, E. H. Friendships and Quarrels Among Preschool Children. Child Dev., 1933, 4: 237–252.
6. HAGMAN, E. P. The Companionship of Preschool Children. Iowa City, Iowa, The University. 1933.
7. JENKENS, G. G. Factors Involved in Children's Friendships. J. Ed. Psychol., 1931, 22: 440–448.
8. PARTEN, M. B. Social Participation Among Preschool Children. J. Abn. and Social Psych., 1932, 27: 243–269.
9. SALUSKY, A. S. Collective Behavior of Children at a Preschool Age. J. Soc. Psychol., 1930, 1: 367–378.
10. VREELAND, F. M., AND COREY, S. M. A Study of College Friendships. J. of Abn. and Soc. Psychol., 1935, 30: 229–236.
11. WELLMAN, B. The School Child's Choice of Companions. J. Ed. Research, 1926, 14: 126–132.
12. WILLIAMS, P. E. A Study of Adolescent Friendship. Ped. Sem., 1923, 30: 340–346.
13. ZALUZHNI, A. S. Formation of Social Habits in Children of Preschool Age. (Vide: Murphy, G., and Murphy, L. B. Experimental Social Psychology, New York, Harper and Bros., 1931, p. 282.)

PROBLEMS FALLING WITHIN THE SCOPE OF PSYCHOLOGICAL COUNSELING[1]

PERCIVAL M. SYMONDS
Teachers College, Columbia University

TRADITIONALLY the scope of the work of the clinical psychologist has been limited. In the laws of the various states a psychologist is recognized as one who certifies to the feeblemindedness of a person in connection with commitment to an institution, or in some states in connection with sterilization proceedings. Actually psychology as a science, with its accumulated wisdom regarding motives and incentives, individual differences, learning, and the nature of personality, provides a basis for professional service of a much more comprehensive character than the measurement of intelligence level by mental tests.

The scope of the services which may reasonably be thought to be within the province of the clinical psychologist can best be approached by a study of the problems coming to the psychologist. *Problem* is defined simply as that which causes complaint and appeal for aid by a person himself or by his associates. A person may complain about his fears, failures, confusions, etc., whereas his associates may complain about his behavior which either annoys or injures another (as criminal behavior) or indicates his failure to manage his own affairs. In order to have these problems more clearly in mind a number of recent lists of the types of problems coming to psychologists for advice, guidance, and help were scanned and a composite inventory of problems was made. The studies which were used in this compilation are listed in the bibliography included at

[1] Paper read at the meeting of the Association of Consulting Psychologists, Vineland, N. J., May, 1936.

the end of this paper. These are inventory studies of problems of children in the pre-school, elementary and secondary school, also at the college and graduate school levels. Of special significance is Ackerson's *Children's Behavior Problems,* Tilson's *Problems of Pre-School Children,* and Wickman's *Children's Behavior and Teachers' Attitudes.*

Each of the problems mentioned in the studies was copied on a separate slip and these were then sorted and re-sorted until like problems were together and a reasonable amount of order appeared from the chaos. The classification of these problems is extremely difficult, owing to the complexity of the organism whose problems they are. It is impossible to present the complete inventory of problems in the limits of this paper, as the typewritten copy covers 34 pages and includes 1071 items. The following outline gives the headings under which the problems were finally organized.

Problems Coming to the Psychological Counselor

- Motor control
  - Nervous tics
  - Coordination
  - Sphincter control
- Speech defects and disorders
- Mental deficiency
- Emotional control
- Fears and anxieties and behavior resulting therefrom
- Habits
  - Health habits—eating, sleeping, etc.
  - Personal care
  - Sex
  - Use of tobacco, alcohol, drugs
  - Money
  - Social relations
    - Adjusting to others
      - Aggressiveness
    - Anti-social, delinquent and criminal tendencies
    - Deception and theft
- Attitudes
  - Toward self
  - Toward other persons
  - Toward natural phenomena

Social problems
  Education
    School adjustment
    Language skills
    Other skills
  Vocation
    Selection
    Success
  Religion
  Marriage and family
  Recreation
  Special problems of adolescence and the involution period

It will be noted that the problems which persons bring to the psychologist are considerably wider in scope than the problems of mental deficiency. All, however, represent failures of learning which result in poor adjustment. It should be recognized, however, that these areas are interrelated and they cannot be thought of simply as so many separate learning inadequacies. Problems of motor control, for instance, may have a physical basis. Bad habits may represent merely lack of education. Poor and faulty habits may be natural protective adjustments toward situations causing fears and anxieties, as well as toward false ideas and attitudes, and it is these mechanisms which the psychiatrist considers his special province. Again these same set of bad habits may be complicated by mental deficiency.

The group of special problems involve psychological issues but are somewhat broader in scope than would normally be included in the training of a psychologist in its limited sense. For instance, to advise concerning family problems one should know something about the sociological and economic aspects of the family, problems of sex, and the servant problem as well as the more strictly psychological problems of personal relationships and child guidance.

The problems in the list of those coming to the attention of the psychologist under the heading of emotional control, fears, and habits parallel in certain respects the list of mental disorders as classified by the American Psychiatric Association.

Classification of Mental Disorders

- Profound personality changes
  - Psychoses
    - Psychogenic
      - Dementia praecox
      - Manic depressive psychoses
      - Involution melancholia
      - Paranoia
    - Organic
      - Infective, exhaustive, and toxic
- Borderline mental disorders
  - Psychoneurosis
    - Hysteria
    - Anxiety neurosis
    - Neurasthenia
    - Compulsion neurosis
    - Traumatic neurosis
  - Constitutional psychopathic states
    - The inadequate personality
    - Emotional instability
    - Pathological lying
    - Kleptomania
    - Sex perversions
    - Drug addiction
- Mental deficiency

They differ with respect to the language used, the problems coming to the attention of psychologists being stated in non-technical language, whereas the psychiatric classification is stated in the customary medical terminology. One must conclude, therefore, that problems coming to the attention of the psychologist differ from the more serious cases of mental disorder only in degree and not in kind.

The real issue of the scope of the psychologist's task lies in the kinds of clinical experience which he receives as well as his technical information and laboratory experience. If the only experience in working with people which a psychology student receives in his period of training is the administration and interpretation of mental tests, then he is qualified to be nothing more than a psychometrician. If his training includes study-

ing and guiding problem children in schools, then with the proper technical background he becomes qualified to become a school psychologist ready to understand and deal with the psychological problems of children in school. I should say that any one whose training did not include internship in a mental hospital was not qualified to deal with the psychoses.

A general rule may be formulated that a psychologist is qualified to deal with any psychological problems including the neuroses and psychoses, provided experience in studying and dealing with such cases is included in the period of training. In order to make clear exactly what this means I have drawn up the following scale of Mental and Behavior Disorders. This scale including as it does on the one hand deviations toward mental disorder leading to the psychoses, and on the other deviations toward behavior disorder leading as one outcome to criminality, is recognized to be complex, but it serves the present purpose. At one extreme individual 1 is described as the perfectly adjusted man while at the other extreme individual 8 has a deep psychosis.

Scale of Mental and Behavior Disorders

Superior Man	1. The person meets the varied demands of life entirely adequately. Life problems are solved satisfactorily. The person is happy himself and is beloved and highly respected by his associates. He faces reality squarely and courageously.
Average Man	2. The person manages his affairs fairly successfully and maintains himself as an independent, self-supporting member of society. Occasionally he makes mistakes in meeting life's problems but they are not serious and are eventually readjusted. He is able to work these problems out himself without assistance. He is fairly satisfied with himself and is respected as an average member of his community. He has the usual mechanisms of rationalization, projection, etc., in meeting rebuffs and thwarting.
Mild Personality Peculiarities	3. The person is still an independent and self-supporting member of society but shows noticeable personality peculiarities. He has difficulty in getting along in school or at work but not to the extent that he is referred for special study or discharged or expelled. He may show signs of feelings of inferiority. He shows the results of faulty training or of ignorance. He has difficulty in adjusting to others. He is recognized as peculiar in one or more respects. These peculiarities, however, are accepted by others as within the

normal range, and he neither seeks advice concerning them nor do others offer him unsolicited advice.

4. The person has mild personality or behavior problems, but his complexes and conflicts do not lead to full-blossomed neuroses. He may show such problems as stuttering, lack of sphincter control, pronounced feelings of inferiority, bad habits such as masturbation, lying, petty stealing. In school he is a problem child and is referred for special study. Is distinctly unhappy at times or gives his associates concern. The difficulty is readily cleared by environmental readjustments or simple psychotherapy.

Adjustable Problems

5. The person has a neurosis which causes concern to the person himself or to his associates or he exhibits psychopathic behavior. He is able to marshal his mental forces so as to control his own affairs, but his difficulty partially incapacitates him, renders him inefficient, or makes him a social problem. On the social side he may be predelinquent, but is still allowed to remain in his school or family. The case requires skilful and persistent psychoanalysis, or extensive environmental readjustments before yielding to treatment.

Neurosis

6. The person has mild psychotic tendencies such as incipient schizophrenia, manic-depressive states, petit mal, etc. Here would be placed the mentally defective children who are placed in special classes or those who are recovering from epidemic encephalitis and require care in special groups. This group would include delinquents who are best handled in special schools. Recovery may be expected if circumstances are favorable, but temporarily the person is removed to a special group or institution as unable to adjust or make normal progress in a regular group as temporarily and mentally incapacitated, or as a menace to society.

Pre-Psychosis

7. The person is psychotic and must be institutionalized or hospitalized. Personality deterioration has gone so far that he is unable to adjust in society or his family. He has lost the power to manage his own affairs. Within the hospital or institution he can take some responsibility, and care for his person, can do simple or skilful tasks, and can be handled in groups.

Psychosis

8. The person has a deep psychosis or is incapacitated by extensive neurological lesions.

Deep Psychosis or Organic Lesions

I should say that psychologists with thorough training in psychology and clinical experience with the type of person with whom they are to deal, but without medical training, should be qualified to deal with individuals desiring study, guidance, and treatment in groups 4 and 5. In many cases, however, the cooperation of a physician is desirable. Every

case in which there are physical symptoms of nervousness, faulty metabolism, anemia, malnutrition, organic impairment of any sort, or physical disease should be referred to a physician for examination. It is a good general rule that treatment for physical malfunctioning or organic impairment should precede or parallel any attempt to secure better mental adjustments.

Groups 5 and 6 do not represent a sharp break. Those in group 6 show progressively greater disorganization and are less able to care for themselves as independent members of society. Psychologically their problems are deeper, less accessible, and less amenable to treatment.

Only those with hospital or institutional experience are qualified to treat individuals showing mental disorder of the levels represented by groups 6, 7, and 8, and this is naturally restricted to those with medical training although there is no reason why a person with psychological training should not give valuable professional service if permitted to gain institutional experience with psychotics. The educational treatment of those mentally deficient children who must be placed in special classes may be the special province of psychologists whose training has specialized in the problems of mental deficiency.

On the other hand, psychiatrists by their training and experience are qualified to handle cases in groups 4 and 5, particularly where it is a matter of emotional conflict. A case that involves problems of school adjustment is probably best handled by a psychologist with experience with school problems, whereas a problem which involves home training may be handled either by a psychologist (provided his training has included problems of this kind) or by a psychiatrist.

Psychologists should, as a matter of professional ethics, be extremely careful not to accept for study or treatment problems which their training does not qualify them to handle. Any case in which there is suspected some neurological disturbance as indicated by tics, tremors, spastic or flaccid reactions, or incoordinations should be referred to a neurologist

for examination. Any case in which the personality disturbance is suspected to have an endocrinological basis should be referred to an endocrinologist for study and endocrinological treatment. Any case in which it is evident that the personality is badly disorganized and disintegrated so that the individual has lost the capacity to manage himself or his affairs should be referred to a psychiatrist or a psychologist whose experience qualifies him to deal with the more severe cases of disorganization. In no case should a psychologist undertake medical treatment or any kind of physical therapy. Cooperation with other professional groups is advisable wherever it seems to be called for. The code of ethics of the Association of Consulting Psychologists clearly states

"Members of this Association will call into consultation the proper specialist, in particular the physician, whenever their own training and experience require such supplementation."

I would even go further and add that in case the training and experience are not adequate the case should not be accepted at all, but should be referred to the proper specialist.

References Containing Inventories of Problems which Are Psychological in Nature

Ackerson, L. *Children's Behavior Problems.* Behavior Research Fund Monographs, vol. 1, Incidence, Genetic and Intellectual Factors. University of Chicago Press, 1931.

Binnewies, W. G. "A Study of the Social, Vocational and Educational Problems of College Girls," *Journal of Educational Sociology,* 5: 82–88, 1931–1932.

Boraas, J. "Troubles of College Freshmen," *School and Society,* 6: 491–494, Oct. 27, 1917.

Bragdon, H. O. *Counseling the College Student.* Harvard University Press, 1929.

Emme, E. E. *The Adjustment Problems of College Freshmen.* Colesby Press, Nashville, Tenn., 1933.

Green, G. "Freshmen Problems," *National Association of Deans of Women Sixteenth Yearbook,* 152–154: 1933.

Hartmann, G. W. "The Classification of Adjustment Problems Among College Students," *Journal of Abnormal & Social Psychology,* 28: 64–69, 1933–1934.

HEATON, K. D. *The Character Emphasis in Education.* Appendix: Bibliography of References on Child Guidance. University of Chicago Press, 1933.

HILDRETH, G. "A Survey of Problem Pupils," *Journal of Educational Research*, 18: 1–14, June, 1928.

LEATHERMAN, Z. E. *A Further Study of the Maladjusted College Student*, Ohio State University Studies, vol. 2, no 2. Ohio State University, 1928.

MEYER, G. "Problems Concerning Children as Reported by Teachers," *California Quarterly of Secondary Education*, 5: 165–170, Jan., 1930.

STOGDILL, R. M. "Attitudes of Parents, Students and Mental Hygienists toward Children's Behavior," *Journal of Social Psychology*, 4: 486–489, 1933.

STOGDILL, R. M. "Parental Attitudes and Mental Hygiene Standards," *Mental Hygiene*, 15: 813–827, 1931.

STRANG, R. "Personal Problems of Students," *Yearbook of the National Association of Deans of Women*, 16: 147–151, 1929.

STRANG, R. *The Role of the Teacher in Personnel Work.* Bureau of Publications, Teachers College, Columbia University, 1932.

STRANG, R. "Problems of Adolescents which Come to Deans," *Junior-Senior High School Clearing House*, 7: 29–34, Sept., 1932.

STRATTON, D. C. *Problems of Students in a Graduate School of Education.* Contributions to Education, No. 550, Teachers College, Columbia University, 1933.

TILSON, M. A. *Problems of Pre-school Children: A Basis for Parental Education.* Contributions to Education, No. 356, Teachers College, Columbia University, 1929.

WICKMAN, E. K. *Children's Behavior and Teachers' Attitudes.* The Commonwealth Fund, Division of Publications, 1928.

AN EVALUATION OF THE KENT ORAL EMERGENCY TEST

MARY I. ELWOOD
Pittsburgh Public Schools

EDWARD M. L. BURCHARD
Torrance State Hospital

FLORENCE M. TEAGARDEN
University of Pittsburgh

IN 1932 Dr. Grace H. Kent published her "Oral Test for Emergency Use in Clinics."[1] The test as published at that time represented some nine years' study and six editions. It was framed expressly for use with adults and was said to require about 10 minutes for administering. The test consists of 25 short questions arranged so as to form three graded scales of 15 questions each, one group of five questions being common to each of the three scales. The Lower Scale was said to be discriminative for mental ages 6 to 10 years; the Middle Scale, 8 to 13; and the Upper Scale 10 to 14. The intention was to present a test which requires no special facilities; makes the minimal demands upon the subject's willingness to respond; is suitable for use in jail, courthouse, or clinic; can be given in a very short time; will indicate further tests of different kinds which should be administered; avoids all use of questions relating to school life or juvenile experiences although use has shown its applicability to children of ages 8 to 13; requires no reading or writing on the part of the testee; does not penalize a speech-defective; avoids questions requiring subjective evaluation; calls for information of a constant and universal nature rather than academic subject matter; and requires only negligible recording.

At the time of the appearance of Dr. Kent's monograph, one of the present writers presented it for discussion at the staff

[1] "Oral Test for Emergency Use in Clinics," *Mental Measurement Monographs*, No. 9, January, 1932, 50 pp.

seminar at the University of Pittsburgh. She and two others in attendance at the seminar determined to try out the scale in an attempt to valuate the claims made for it. It seemed impossible that an oral test requiring no more than 10 minutes could give much of diagnostic value.

The findings reported here include only cases upon whom a Kent Emergency Test and a Stanford Revision of the Binet Test were administered at the same time. In all but a very few cases an individual's ratings on the Kent and on the Binet were secured by the same examiner. Dr. Mary I. Elwood, of the City Schools of Pittsburgh, used one scale of the Kent on 215 children to whom she was also giving Binets. These children for the most part are those who are being considered for admission to opportunity rooms or special classes. Only 34 of the 215 children have Binet IQ's of 100 or more. Dr. Edward M. L. Burchard of the staff of the Torrance State Hospital examined 25 patients using, along with other tests, both the Kent and the Binet. With over half of these cases he used two scales of the Kent and with a little less than half he used three scales. It will be noted in the table below that some of his cases are children. These were children examined in Centers that are served by the State Hospital Traveling Clinic. Dr. Florence M. Teagarden of the Psychological Clinic of the University of Pittsburgh, gave Kents to 85 patients to whom Binets were given at the same time. These individuals represent the usual run-of-mine cases referred to a psychological clinic—behavior problems, cases referred by social agencies, and the like. About 47 per cent of these Clinic cases were given two scales; 27 per cent, three scales; and 26 per cent, one scale. The accompanying table shows something of the composition of the three groups of cases.

As one looks over Table I, a question comes to mind. Since 14+ is the highest mental age obtainable on the Kent, is it possible that some of those individuals having Binet mental ages of over 14 are penalized on the Kent by hitting the top? Scrutiny of actual records shows that this did not occur in a

TABLE I

325 Cases to Whom both Binets and Kents Were Administered

	NUMBER OF CASES	C.A. RANGE	BINET M.A. RANGE	KENT M.A. RANGE
State Hospital	25	10–8 to 67 yrs.	5– 0 to 11–5	5–4 to 12–5
City Schools	215	6–0 to 15–2 yrs.	5–11 to 17–6	5–1 to 14–8
Psychological Clinic	85	7–1 to 50 yrs.	4–10 to 16–5	6–1 to 14–10

single instance in the 325 cases. There were a few records in which testees got scores above the highest norms on the Kent, but examination shows that they happened not to be ones whose Binet mental ages were over 14 or 14½. All cases which ran either above or below Kent norms were discarded for purposes of the present study.

When the combined cases of the three investigators neared the 300 mark, the records were turned over to the one who had presented the Kent monograph originally. The statistical treatment from that point on is solely her responsibility. It may be questioned, but at least it makes possible the comparison which we had desired to make. That is to say, we wished to get an idea of how closely parallel the Binet and Kent mental ages were running.

In the Kent monograph there is no interpolation made within years. Any score from 13 through 17, for example, on the Low Scale gives a mental age equivalent of seven years; while any score from 18 through 22 gives a mental age equivalent of eight years. Similarly on the other two scales a subject rates a mental age of nine, or 10, or 11, or whatever the case may be but there is no provision made for discriminating *within* any one of these years. In order to make possible comparison with Binet mental ages it seemed necessary to interpolate Kent scores. In some instances on the Kent three scale-scores are assigned to one mental age equivalent; in some, four; in some, five; and in some, six. It was a little difficult to determine the best method of weighting these varying amounts of year credit. It is perhaps not necessary to go into detail here as to how it was finally

done. Other workers might well use different interpolation methods, but it is doubtful whether comparison-results would differ much from one method to another. Suffice it to say that by the method adopted in the present study it was possible to get Kent mental ages of seven years, one month; seven years, four months; eight years, one month; eight years, four months; etc.

After having interpolated Kent scores into these mental age equivalents, Kent IQ's were then found. This was done because it was easier for correlation purposes to have only one number to handle (IQ 87 compared with IQ 90, for example) than two numbers (M.A. seven years, four months compared with M.A. seven years, nine months). There would seem to be ample precedent for such a procedure. It was then possible to compare Kent IQ's and Binet IQ's. It was also possible to compare IQ's found by the use of different Kent scales where more than one scale had been used with the same individual. The correlations are given in Table II.

TABLE II

Product Moment Correlations Between Kent IQ's and Binet IQ's

TESTS	NUMBER OF CASES	PEARSON r	P.E.
Kent Low and Binet	252	.88	.0096
Kent Middle and Binet	125	.90	.0115
Kent Upper and Binet	61	.86	.0225
Kent Low and Kent Middle	57	.96	.007
Kent Low and Kent Upper	33	.91	.0202
Kent Middle and Kent Upper	58	.89	.0018

Even those of us who had been using the Kent and finding that it rather consistently gave a close approximation to the Binet rating were surprised at the size of the correlations. They are as high as many reported in the literature where justification of a test is being sought.

In order to make another comparison it was decided to count the number of cases in which the mental age ratings were exactly the same on the Binet and on the Kent and the number

of cases having varying degrees of divergence on the Kent above or below the Binet. Because the range of divergences is so great the differences are placed in class intervals of four months. These differences are shown in the following table.

An analysis of the figures in Table III introduces interesting speculations. Let us look at the distribution for the Low Scale of the Kent. Fifty per cent of the cases in this distribution lie within the range represented by the deviations "four months below" and "eight months above." Every case falling at or above "24 months above" is a State Hospital case. One can only conjecture whether these six cases who deviate so extremely do so because of the condition which referred them to a State Hospital.

In the part of Table III showing the distribution for the Middle Scale of the Kent, 50 per cent of the cases lie within the range represented by "four months below" and "12 months above." Every case falling at or above "28 months above" is a case belonging to the State Hospital group. Of the six cases falling below "20 months below," two are State Hospital cases, three are Psychological Clinic cases referred because of peculiar behavior not explainable on the basis of intelligence, and one case belongs to the Public School group and is not at this time identifiable.

In the Upper Kent Scale part of Table III, 50 per cent of the cases lie within the range indicated by "16 months below" and "12 months above." Of the eight cases lying at or above "20 months above," seven are State Hospital cases. The other case is a 33-year-old Psychological Clinic patient who had recently in a fit of anger run her fist through a window pane, on another occasion had hit her mother over the head with a broomstick with considerable violence, had injured guests in the home, and when not in fits of anger sat with a mask-like expression of extreme depression. Of the six cases lying below "24 months below," four are Psychological Clinic cases who are either extreme behavior deviates or who have recently undergone severe emotional experiences. The other two cases

TABLE III

Number of Cases Deviating in Kent M.A. from Binet M.A. by Specified Number of Months

KENT LESS THAN BINET											
48 mos.	44 mos.	40 mos.	36 mos.	32 mos.	28 mos.	24 mos.	20 mos.	16 mos.	12 mos.	8 mos.	4 mos
						1	3	6	14	35	29
Low Scale Number = 252											
	1	1	2	0	2	0	3	5	5	10	10
Middle Scale Number = 125											
1	0	2	3	0	0	0	7	3	2	6	4
Upper Scale Number = 61											

TABLE III—(*Continued*)

KENT MORE THAN BINET										
SAME	4 mos.	8 mos.	12 mos.	16 mos.	20 mos.	24 mos.	28 mos.	32 mos.	36 mos.	40 mos.
15	44	46	23	18	11	2	2	0	0	2
Low Scale Number = 252										
3	23	22	14	9	4	5	2	1	2	1
Middle Scale Number = 125										
1	3	9	8	4	1	3	2	1	1	
Upper Scale Number = 61										

belong to the Public School group. One of these individuals is a 10-year-old girl with a Binet IQ of 156. Her Kent score was obtained on the Upper Scale and therefore probably calls for information outside the experience of even a very superior 10-year-old. The other case is that of a boy with a Binet IQ of 150. He is living in a most unusual family situation which is far from ideal from a mental hygiene point of view.

To recapitulate, it appears (1) that on each of the three Kent scales the middle 50 per cent of cases lie within a range of test-

result deviations no greater than that usually found in comparing results obtained on different tests. Correlations run somewhat higher than many test-result comparisons reported in the literature. (2) While the number of cases where there is extreme disparity between test results is not large enough to warrant final conclusions, suggestive trends may be indicated. Of the records studied here it is the State Hospital cases that contribute most to the number that run very much higher on Kent than on Binet. If on 1,000 cases this should prove to be true also, some valuable light might be thrown upon the relative effect of various mental derangements upon the functions represented in the Binet and in the Kent. (3) Of the cases within our study who fall materially lower on Kent than on Binet a disproportionately large number are Psychological Clinic cases referred for behavior disorders not explainable on the basis of intelligence. What a study of 1000 similar cases would show cannot even be conjectured but it is to be hoped that eventually sufficient data will be accumulated to make such investigation possible.

Perhaps a few short examples will illustrate some of the values of the Kent and at the same time show that its use is by no means a single short-cut toward diagnosis or evaluation. Let us take the case of a widower of 55 referred to the Psychological Clinic by the Juvenile Court because there was a question as to the suitability of the home for the rearing of his nine children. Mr. Q. appeared at the Clinic not knowing exactly what the trip was all about but quite willing to go along with whatever was expected of him. It happened, however, that he had come without his glasses which made certain omissions on the Binet necessary. Mr. Q. had not been in school for years and his coal-mining occupation had made no demands of a literary or academic nature. He had, however, taken some correspondence school courses and had done some pencil sketching. Much of the Binet seemed rather unfair for the examination of such a man. His approximate mental age (allowing for omissions noted above) was 14 years and three months. It

was decided to try the Kent since it represents more the type of experience that such a man would have had. On the Middle Scale he secured by our interpolation a mental age of 14 years and 10 months. On the Upper Scale he "ran out at the top" making the highest score possible, which would give him a mental age of "more than 14." Whether this example tends to vindicate the use of the Binet for such cases in spite of apparent inappropriatenesses or whether it indicates that Binet and Kent probably measure the same thing with approximately the same results one does not know. However that may be, one felt that an attempt to answer the question raised by the Juvenile Court could proceed on the assumption of a mental age of not far from 14 years.

Mr. A. was a 45-year-old Italian who can neither read nor write English. His Binet vocabulary scored slightly below the 10-year level and his mental age was 10 years, 8 months. It was felt that he had probably been somewhat penalized on the Binet because of his language handicap. Furthermore, Mr. A had had an excellent work record. For years he had sold candy, sandwiches, and ice cream on excursion trains. He had made enough money to buy the lot on which his little house stood, had built a larger brick house on the front of the lot, and had rented out the little house. All this made one wonder whether his intelligence might not actually be higher than the Binet indicated. The Kent was administered. By our interpolation he made on the Low Scale a mental age of 10 years, 7 months; on the Middle Scale, 10 years, 10 months; and on the Upper Scale, 11 years, 4 months. The Binet results were apparently largely corroborated and there was added evidence that Mr. A. was not an intellectual giant in spite of his apparent economic acumen. Later information shows that the building of the new house had included the installation of a $200 chandelier in a house worth probably $4000 or $5000. He had also allowed the rear house to fall into ruins. Rumors suggest that maybe part of his income has been from sources other than his train work and smacked suggestively of gam-

bling. One wonders whether this man is not in all probability either a borderline or perhaps a slightly feebleminded individual who had a very lucky vocational break by getting into an occupation where a routine "line" of English easily suffices; where making simple change, regularity and affability are about the only requisites. Would that we were able so adequately to place all such borderline cases!

Mrs. D. is 50 years of age, the record says. She knows neither her age nor her birthday. Her husband was recently convicted of murder and was electrocuted. She can neither read nor write. When she attempted to "draw a diamond with a pen" at the Binet seven-year level she caught the pen too near the end. She was so unaccustomed to holding a pen that instead of slipping the fingers of her right hand along toward the pen point, she reached over with her left hand and pulled the pen into position. Her basal age on the Binet was at the five-year level, and she had no successes beyond seven years. Her mental age was 6 years, 6 months. For such a woman there certainly is considerable question as to how many tests are necessary in order to get an adequate estimate of intelligence and how one may avoid penalizing for lack of opportunity. On the Low Kent Scale she made, by our interpolation, a mental age of 7 years, 10 months. On the Middle Scale her score was too low to compare with Middle Scale norms; in other words, it was below eight years. Matters of common observation such as what money is made of, how many minutes there are in an hour, why it is colder at night than it is in the daytime, the names of some large cities, and the time of day at which one's shadow is the shortest were just as foreign to her as the more linguistic and academic type of Binet material.

In conclusion, it may be said that on over 300 cases here considered surprisingly high correlations were found between results on the Kent Emergency Test, which takes little more than 10 minutes to administer, and the results on the Binet test. In the second place, for many of these 325 cases it was felt that

the added information gained by the use of the Kent was extremely valuable in getting a more nearly real picture of the testee. In the third place, it would appear that large discrepancies between Binet results and Kent results may prove to occur most frequently in psychotic cases. Further study of this point would seem to be justified. In the fourth place, the present authors wish to concur with Dr. Kent that the Emergency Test should be used not instead of but as a supplement to other test measures.

A STUDY OF THE EFFECT OF INDUCED MANUAL HANDICAPS ON AUTOMOTIVE PERFORMANCE IN RELATION TO REACTION TIME

LeVERNE JOHNSON AND A. R. LAUER
Iowa State College

STATEMENT OF THE PROBLEM

THE question often arises as to how much one's ability to drive an automobile is affected by the loss of one arm or leg. A one-arm driver is assumed to be unsafe on the highway. From the practical as well as the scientific point of view it is important to know how much the risk increases with the loss of a member. Will a person using one arm compensate by driving slower and by taking greater precaution to the extent that he will offset the loss of an arm? How much less efficient is a man driving with one arm than one driving with both? Does the driver with high reaction time tend to compensate more than one with low reaction time?

With these questions in mind the present study was set for experimental investigation.

APPARATUS AND PROCEDURE

The standardized indoor driving test described by Lauer and Kotvis (1934) was used for the right-left hand driving performance. It consists of a miniature car driven over a roadway laid out on an endless leather belt. The small car is guided by the subject from full-sized standard car controls. When corrected for attenuation, the validity of the test is about +.62 using supervisors' ratings as the criterion. The reliability has been found to be about +.85 for an unselected group of subjects.

Manipulative ability is measured by the number of contacts on the traveling roadway hit by the miniature car. A record of contacts is made by an automatic electric counter. The time for a record is measured and recorded by a counter operated from a synchronous motor and a circuit breaker. Such "time" is the duration of seven revolutions of the roadway, the speed of which is controlled by the driver. Contacts are recorded for the same period. These self-recording instruments eliminate all subjectivity. Errors, failures to stop for stop signs, failure to follow directions and road signs, etc., are recorded by the experimenter who sits on the seat beside the subject.

SERIES I

Fifty-three subjects were used in this study, forty-nine of whom were undergraduates who volunteered from various psychology classes. Three were graduate students and one was an instructor. The three graduate students and the instructor had driven the car before, but not with only one hand. It is conceded that college students may be aware of experimental conditions and consequently not be a random sample of all automobile drivers, but this should not affect their comparative scores.

Directions for the experiment were typewritten on a card and read by the subject before the test. The only verbal directions given were instructions as to which hand to use first and for a change of hands. These directions were not printed, as a different order was followed for each succeeding subject in order to minimize any systematic error due to practice effect or ennui. Each subject was given a preliminary trial to familiarize him with the procedure before the experimental series began. The order for the experimental procedure is given below:

Subject 1. Right—Left—Both
Subject 2. Left—Both—Right
Subject 3. Right—Both—Left
Subject 4. Same as 1, etc.

The complete test of three runs took from ten to fifteen minutes for each subject. Scarcely any time was needed between trials, as all data were recorded as taken.

The reaction time measurements for this study were taken by means of a mechanical chronoscope developed by Allgaier (1935). An auditory stimulus was used in all cases, and the subject was required to depress the reaction key approximately 2 mm. against a resistance of 50 grams. Forty-five subjects were used, thirty-five of whom had also taken the driving test. The latter subjects were used for a comparison of slow and rapid reaction time in relation to compensation for a physical handicap.

Each subject was given thirty trials with no preliminary period. The reliability coefficient was determined by correlating the first fifteen trials with the last fifteen of the forty-five original subjects. It was found to be +.56 which is significant according to Fisher (1930). Corrected by the Spearman-Brown formula the reliability was found to be +.78 for the mean of thirty reactions.

For reaction time both "time" scores and the average deviation were used to correlate with the driving score. The sum of the three trials was used for the latter.

Correlations were calculated by the Pearson Product-Moment Method, using a formula which was written by Wallace and Snedecor (1931), as a modification of the earlier formula by Ayres:

$$r = \frac{\text{Sum } AX - \text{Sum } A(M_x)}{\sqrt{\text{Sum } A^2 - \text{Sum } A(M_x) \cdot \text{Sum } X^2 - \text{Sum } X(M_x)}}$$

RESULTS AND CONCLUSIONS

As far as manipulative ability is concerned Table I shows that there are very slight differences between manipulative scores of this type using both hands and either hand alone. The average number of contacts hit using both hands was only 2.44 more than the right hand score and only 1.98 more than the left hand score. Between the left and right hands there

TABLE I

Contacts or Manipulative Scores

Means	B = 25.80	R = 23.40	L = 23.90
Standard deviation	B = 10.90	R = 7.10	L = 6.70
Standard error of mean	B = 1.50	R = 0.97	L = 0.93
Differences of means	B – R = 2.44	B – L = 1.98	
Standard error of difference	B & R = 1.43	B & L = 1.41	
Critical Ratio	B & R = 1.70	B & L = 1.40	
	B = Both	R = Right	L = Left

was practically no difference although the left hand score was a little more consistent than either of the other scores, its standard deviation being only 6.7 as compared to 7.1 and 10.9 respectively.

The correlation between the right hand and both hands was +.44 and between the left and both +.47. Both of these correlations are significant but indicate low reliability for the short test with the subjects used. This reliability is considered by Kelly (1924) to be adequate for group comparisons if the range of the traits are limited.

In working out the group differences the critical ratios were 1.7 and 1.4, neither being significant. In other words the drop in efficiency due to the use of one arm is very small. In fact the right or left hand alone is about 90 per cent as efficient as both hands.

TABLE II

Time or Speed of Completing the Test*

	(Controlled by the subject)		
Means	B = 78.80	R = 85.60	L = 85.80
Standard deviations	B = 15.80	R = 15.90	L = 15.80
Standard error or means	B = 2.17	R = 2.18	L = 2.17
Differences of means	R – B = 6.74	L – B = 6.98	
Standard error of differences	B & R = 3.07	B & L = 3.06	
Critical ratios	B & R = 2.18	B & L = 2.70	
	B = Both	R = Right	L = Left

* Time units are equal to two seconds each, but were used directly as recorded for comparative values.

Here the difference between using one hand and both hands is greater than in the case of manipulative scores, indicating that the driver using one arm tends to compensate for this handicap by slowing down his speed. The figures show that the person using one arm "drives" about 8 per cent slower than when using two arms. The difference is significant, the critical values being verified by Fisher's table of "t" values.

Assuming that a one-arm driver does drive slower over the highway, as on this apparatus, the fact may be interpreted in two ways. It may be taken as a loss in efficiency or as an indication that the driver realizes his handicap and takes greater precautions because of it. The results shown in table I on manipulative scores tend to substantiate the latter view, for it was shown that there is little loss in efficiency of manipulation due to the use of one member only. A correlation of +.72 was obtained between both hands and the right, and +.73 between both and the left. Correcting by the Spearman-Brown formula the reliability was +.84, which is quite satisfactory when the scores are added together as a composite.

TABLE III

Errors = Observational Ability During the Test

Means	B = 2.10	R = 2.30	L = 2.20
Standard deviations	B = 1.50	R = 1.60	L = 1.30
Standard error of means	B = 0.21	R = 0.22	L = 0.18
Differences of means	R − B = .21		L − B = .13
Standard error of differences	B & R = .30		B & L = .24
Critical ratios	B & R = .70		B & L = .54
	B = Both	R = Right	L = Left

It is quite clear that errors: stop signs, road markers and other rules disobeyed are very slightly affected by a handicap in manipulative ability. It might be expected that a driver could see a stop sign as well when driving with one hand as with two except for the direction of his attention. By slowing up he apparently compensates for any such possible influence. There is no significant difference between the number of errors

made with either or both hands. The relatively small advantage shown for both hands is probably due to the greater concentration upon manipulation when driving with one hand. Thus, with the attention directed more to manipulation, the driver would be less efficient in observation.

The correlation of errors with manipulative ability was found to be – .21 which is only barely significant. Considering the direction of score values this would be interpreted as a slight indication that the person who manipulates better tends to have better observation and vice versa.

SERIES II

In the use of reaction for predicting manipulative ability of the type described in Series I, it was found that the average deviation of the reaction time was probably more valuable than the over-all time. Over-all reaction time (sum of the 30 trials) gave the following correlations with the respective measures on the driving apparatus:

Contacts	– .14
Errors	.00
Time	+ .11

The P. E. of r is about $\pm$ .12.

None of these correlations are significant, indicating that over-all reaction time is of little value as a criterion of manipulative ability as measured by the apparatus described. The correlation with "contacts" after correction for attenuation was – .25, which is also low. The negative correlation is a function of the mode of writing the score only.

Average deviation score on reaction time gave the following correlations with the same measures as described above:

Contacts	– .19
Errors	+ .10
Time	+ .03

The P. E. of r is about $\pm$ .10.

None of these correlations are highly significant. The correlation with contacts in this case after correction for attenuation

was – .32 which is barely significant. The correlation between average deviation and over-all reaction time was + .68 ± .08, indicating a fairly close relationship between the two. This correlation is in a measure spurious, since the values tend to vary together.

A multiple correlation of $R_{ca \cdot r}$, contacts estimated from average deviations and over-all reaction time is + .22, which is only slightly higher than the zero order correlation, r_{ac}, – .1850. A multiple regression equation of these three variables shows that average deviation is four times as valuable as over-all reaction time in this case.

The reliability of the over-all reaction time was + .56 while the reliability of the average deviation score was + .63, which further substantiates the relative importance of the two measures in estimating motor performance.

TABLE IV

Average Percentage Loss in Efficiency Due to the Use of Only One Arm

Low deviation group	9.44	Slow reaction time group	8.17
High " "	10.56	Fast " " "	13.46

These figures show one outstanding fact. The group having the faster reaction time was affected more by the handicap than the slow group. Also the group having the higher average deviation was somewhat more affected than the group with the low deviation. The latter difference, however, was not significant when comparisons were made by analysis of variance.

Considering the entire group, it seems that those with the slow reaction time and low deviation, the "slow but steady," are less affected by a manual handicap of the type prescribed in this study. A possible explanation for this fact is that the fast reaction-time group may be more erratic and this tendency carries over to driving manipulations. Also the fact that the fast group was slightly more efficient in the beginning and therefore had more chance of loss, adds to the reason for the difference. However, the small differences in their initial

scores cannot account for the large differences in the loss of efficiency.

SUMMARY AND CONCLUSIONS

1. There is little difference between the manipulative ability of both hands as compared with either hand alone on apparatus of the type used in this experiment.

2. The right or left hand alone is about 90 per cent as efficient as when both hands are used together.

3. The subject using one hand alone takes about 8 per cent longer time than when using both hands.

4. Errors (including stop signs, road markers, etc., missed) are very slightly affected by a handicap in manipulative ability; that is, there is no significant difference between the number of errors made with either or both hands under the conditions of this experiment.

5. Average deviation is more valuable than over-all reaction time in estimating performance of the type measured. By weighting according to "beta" values the average deviation score is approximately four times as important as over-all reaction time score.

6. The multiple correlation of manipulative ability estimated from average deviation and reaction time was found to be + .2222.

7. All correlations between reaction time and driving performance were low and insignificant. Some were slightly significant when corrected for attenuation. The latter procedure seems justified when actual relations are desired rather than pragmatic values.

8. Subjects with short reaction time were more affected by restricting one arm than those with slow reaction time.

9. Those with low average deviation scores were less affected by the use of but one arm than those with high average deviation.

10. These conclusions are offered insofar as they pertain to the experimental data presented. Further experimentation is needed before final conclusions are offered. They tend to con-

firm earlier studies which indicate that driving performance is a complex function and cannot be at all adequately determined by any single measure of isolated response patterns known to be required in that performance. The relation of any single variable to driving performance is conditioned by various and sundry external or internal factors which may constantly change. It is necessary to know many things about an automobile driver to be able to predict accurately his liability to accident insofar as such is related to his manipulative ability.

REFERENCES

1. Allgaier, Earl. A Portable Chronoscope. 1935. Amer. Jour. of Psychology.
2. Fisher, R. A. Statistical Methods for Research Workers. 1930. Oliver and Boyd.
3. Kelley, T. L. Statistical Method. 1924. Macmillan and Company, 221.
4. Lauer, A. R. and Kotvis, H. L. Automobile Manipulation in Relation to Vision. 1934. Journal of Applied Psychology, *18*, no. 3; 422–31.
5. Wallace, H. A. and Snedecor, G. W. Correlation and Machine Calculation. 1931. Iowa State College Bulletin, *30*, no. 4; 8.

PSYCHOLOGICAL EFFECTS OF CHANGING THE FORM OF A TRADE NAME

ALBERT B. BLANKENSHIP AND HOWARD R. TAYLOR
University of Oregon

I. INTRODUCTION

THE Jantzen Knitting Mills, makers of the Jantzen Swimming Suit, considered a change of form of the word "Jantzen" in all of their advertisements from the script to the block form, in order to avoid asymmetry and waste of space. It was feared that the change might affect the goodwill built up in connection with the well-known script form of the company's name. Mr. Carl Vreeland, of the Jantzen Mills, suggested an experimental study.

The following experimental study was designed to answer three questions: (1) does the change in print affect association of product with the trade name, (2) does the change affect confidence in genuineness of the product, and (3) is there any preference for one style of print over the other in a typical advertisement?

The method of association is not a new one in advertising analysis.[1] Likewise, work has been done on trade-mark infringement,[2] but no work has been done to measure change in confidence for the same trade-name appearing in different types.

[1] For a historical survey of the method of association used in advertising analysis, see Joseph L. Holmes, "The Free Association Method as a Measure of the Efficiency of Advertising," JOUR. APPL. PSYCHOL., 1925, 9, 60–65.

[2] One of the earliest studies along this line from the psychological point of view is reported by Richard H. Paytner, Jr., "A Psychological Investigation of the Likelihood of Confusion between the Words 'Coca-Cola' and 'Chero-Cola,'" JOUR. APPL. PSYCHOL., 1919, 3, 329–351.

Investigations have been made on preferences for styles of type.[3] It is commonly acknowledged that a thin, fine style of type creates a better impression than a thick, heavy type. In most studies of this sort the "order of merit" method has been used.[4]

The reactions of a group of 92 subjects were used for the first two parts of this experiment. In the third part of the experiment, 75 subjects, a totally different group, were used. All subjects were undergraduate students in psychology classes at the University of Oregon.

II. THE FIRST PART OF THE EXPERIMENT

The subjects were divided into two groups. The 47 subjects in the first group were shown a list of five trade names, presented successively on a screen, all of which appeared in script form. The second group, 45 subjects, were shown the same list in block form. Excluding the word "Jantzen," two forms in each list were "correct," judged by the form in which they usually appeared in advertisements. The names were "Philco," "Whitman's," "Jantzen," "Coca-Cola," and "Texaco," in that order. The experimenters felt justified in not varying the order of presentation, as each subject was required to respond to each word. The subjects were told, in standardized fashion, that they were to indicate, on sheets of paper, the kind of product each name made them think of. It will immediately be observed, as indicated in Table I, that in all cases the percentage of individuals reporting the correct associations was quite high, and that the form of the word had little effect upon this product association.

Table II presents statistical evidence that the percentage of correct associations is not significantly affected by the form of the trade name.

[3] Probably the most inclusive study in this field is that by A. T. Poffenberger and R. B. Franken, "A Study of the Appropriateness of Type Faces," Jour. Appl. Psychol., 1923, 7, 312–329.

[4] The first study using this method was done by E. K. Strong, "Application of Order of Merit Method to Advertising," *Jour. Phil., Psychol., etc.*, 1911, 8, 600–605.

TABLE I

Results of the First Part of the Experiment Showing Per Cent of Correct Associations

STIMULUS WORD	PER CENT IN GROUP ONE (SCRIPT)	PER CENT IN GROUP TWO (BLOCK)
Philco	95.7	100.
Whitman's	97.9	97.4
Jantzen	97.9	100.
Coca-Cola	100.	100.
Texaco	100.	100.

TABLE II

Difference of the Percentage Correct and Probable Errors of the Difference for the Data in Table I

CONDITIONS COMPARED	DIFF. OF PER CENT COMPARED	P.E. (DIFF.)
Philco script-block	4.3	1.93
Whitman's script-block	.6	2.06
Jantzen script-block	2.1	1.41
Coca-Cola script-block	0.0	.....
Texaco script-block	0.0	.....

It may be concluded that change in the form of trade-name has little effect upon product association with the name.

III. THE SECOND PART OF THE EXPERIMENT

The second part of the experiment, designed to answer the second question set, made use of the same group of subjects, divided in the same manner. Each group was shown both script form and block form of each of the same trade names referred to, in the same order. The subjects were told in standardized fashion to record, in each case, which form of the trade-name they had more confidence in, so far as genuineness was concerned. Table III shows quite clearly that there were large differences in confidence for the two groups only in the case of the Philco and Texaco names.

Table IV shows the difference of per cent correct choices, and the probable errors of these differences. From this table it will be noted that both of the differences in question are statistically significant. This would seem to indicate that when the

TABLE III

Results of the Second Part of the Experiment Showing Per Cent of Confidence in Names of Each Form

STIMULUS	PER CENT IN GROUP ONE	PER CENT IN GROUP TWO
Philco script	95.7	48.9
Philco block (correct)	4.3	51.1
Whitman's script (correct)	95.7	88.9
Whitman's block	4.3	11.1
Jantzen script (old form)	95.7	95.6
Jantzen block (proposed form)	4.3	4.4
Coca-Cola script (correct)	100.	97.7
Coca-Cola block	0.	2.3
Texaco script	25.5	0.
Texaco block (correct)	74.5	100.

TABLE IV

Difference of Per Cent Correct Choices and Probable Errors of the Differences for the Data in Table III

CONDITIONS COMPARED	DIFF. IN PER CENT CONFIDENCE	P.E. (DIFF.)
Philco choice for two groups	46.8	2.66
Whitman's choice for two groups	6.8	2.27
Jantzen choice for two groups	.1	2.01
Coca-Cola choice for two groups	2.3	1.45
Texaco choice for two groups	25.5	4.29

individual was not sure of which of the two forms of a name was genuine, he tended to be influenced by the form of the name he had been shown in the first part of the experiment. Group one, for example, had been shown the script forms of the names in the first part of the experiment; in the second part of the experiment this group chose the script form of the names "Philco" and "Texaco" significantly more than did the other group, who had been shown the block forms in the first part of the experiment.

There are no significant differences between the choices of the two groups for the Whitman's, Jantzen, or Coca-Cola names. Both groups were apparently quite sure of the genuine form of each of these names. Our particular interest, of

course, is the choice of each of the groups for the script, or familiar form of Jantzen, as the genuine.

IV. THE THIRD PART OF THE EXPERIMENT

In the third part of the experiment, designed to answer the third question stated at the beginning of this paper, a different group of subjects was used so that attention would not be called to the form of the name. Two Jantzen advertisements, identical save for the form of the name, were presented successively on the screen. The order of presentation was counterbalanced for the two groups, so that each advertisement might have equal average position. The subjects were asked to record their preference for one advertisement or the other, and the reason for their choice. If they had no preference, they were to so record.

It was pointed out to the investigators that merely because the subjects did not recognize the fact that the type used in the word "Jantzen" was different in the two advertisements, that fact did not prove that the subjects were not reacting to the difference. Hence the experimenters took particular care to note whether one form of the advertisement was consistently preferred over the other in each group into which the subjects were divided (for purposes of counterbalancing position of the advertisements). Table V indicates that in the first group,

TABLE V

Preference for Each Form, in Per Cent for Each Group

CHOICE	GROUP ONE (ORDER : SCRIPT, BLOCK)			GROUP TWO (ORDER : BLOCK, SCRIPT)		
	Total Per Cent	Due to Type	Other Reason	Total Per Cent	Due to Type	Other Reason
Script	14.3	2.4	11.9	42.4	3.	39.4
Block	21.3	2.4	19.	21.2	0.	21.2
None	64.3	—	—	36.4	—	—

where the order of presentation was first the script and then the block, that the block form was preferred more often than

the script, though the majority of subjects had no choice. For the group of subjects who were presented the advertisements in the order of block and then script, a few more preferred the script form! And in very few cases was the difference in type observed.

If the figures for the two groups be combined, Table VI is secured. This table shows some significant facts. In the first

TABLE VI

Preference for Each Form, in Per Cent, for Total Group

CHOICE	TOTAL PER CENT	PER CENT DUE TO TYPE	PER CENT GIVING OTHER REASON
Script	26.7	2.7	24.
Block	21.3	1.3	20.
None	52.	—	.....

place, only 4 per cent of the total group of subjects observed the change. The majority of subjects had no preference for either of the two advertisements. The data would seem to indicate that there were no differences in affective value of the two forms for subjects who did not specifically observe the difference. Hence it might be concluded that the affective values of the two forms were different for different subjects, but cancelled out in the whole group. However, when results were tabulated according to the order of presentation of the advertisements, different conclusions were indicated. Table VII shows that the number having no preference at all could

TABLE VII

Choices According to Order of Presentation

PREFERENCE	PER CENT IN GROUP ONE	PER CENT IN GROUP TWO	TOTAL PER CENT
1st advertisement	11.9	21.2	16
2nd advertisement	19.	39.4	28

be increased if one could eliminate, in some manner, the presentation of two successive stimuli without calling attention to the difference in type in the word "Jantzen." The table shows that the second advertisement, regardless of the kind of type

used, was preferred by those who gave other reasons than type differences as the cause of their choice. This would indicate that the order of presentation was involved, rather than conscious or subconscious observation of type differences.[5] Additional evidence points towards the conclusion that type differences were not noted, in general. Most of the subjects quite apparently focused their attention on the human figures in the advertisements, rather than on the printed matter.[6] Most of the Jantzen advertisements, of course, openly make use of "sex appeal." When the advertisements containing figures of girls attired in bathing suits were flashed on the screen before a group of subjects, about half of whom were boys, there was loud whistling, indicating that the girls were the center of attention in the advertisements.

V. SUMMARY AND CONCLUSIONS

The experimenters were interested in investigating psychological effects of changing the form of the word "Jantzen" from the well-known script to a block form. The experiment herein reported was designed to answer three questions: (1) does the change of type affect product-association with the name, (2) does the change in form affect confidence in genuineness of the product, and (3) is there any preference for one style of print over the other in a typical advertisement? The following conclusions were drawn as a result of the experiment:

[5] This phenomenon may likely be explained in terms of attention. The subjects probably used the fading first exposure as a basis of comparison for the second, which, in contrast, appeared brighter and clearer, explaining the preference for the second. Actually, many subjects did report their preference as being due to the brightness or clearness of the second advertisement, regardless of the order of presentation. It is important to note, however, that those choosing the first advertisement for other reasons than type differences in most cases recognized the basis for their choice, order of presentation.

[6] Nixon, in a recent article, concluded that pictures of humans are one of the greatest attention-getters in advertisements. H. K. Nixon, "Two Studies of Attention to Advertisements," Jour. Appl. Psychol., 1925, 9, 176–188.

1. The change in the form of the name has little, if any, effect on the product associated with that name.

2. When a direct comparison is made between the two forms of the name, the subjects choose the familiar (script) form as being the better guarantee of genuineness.

3. The change in the form of the word has very little effect affectively when used in the situation of the advertisement. A majority of the subjects show no preference for either advertisement, and only 4 per cent of the group noted the difference.

Other interesting observations were made, not directly in line with the purposes of the experiment:

1. When the subject has any doubt as to the correct form of a trade-name, he is influenced by what form the experimenter has previously shown him. When the subject is fairly certain as to the correct form, he is not influenced by what the experimenter has already shown him.

2. The human figure is the center of attention when it is used in advertisements.

3. When two almost identical advertisements are presented successively, many subjects will choose the second one presented, due entirely to order of presentation. The order is not recognized as the basis for choice.

4. When subjects prefer the first advertisement in the presentation, they generally realize that the preference is due to the order.

THE RELIABILITY OF CERTAIN GROUP INTELLIGENCE TESTS

F. H. FINCH
State Department of Education
AND
M. E. ODOROFF
Division of Research, State Board of Control, St. Paul, Minnesota

THE testing movement has been characterized throughout its history by an enormous expenditure of effort directed toward the construction of new instruments. There has been, on the other hand, relatively little research designed to discover the comparative value of the numerous measuring devices from which one might choose when confronted with any of the wide variety of problems that call for the application of psychological tests. One of the most important criteria to be considered in the choice of any test is obviously that of reliability. The reliability of a typical test can be determined in a fairly direct and simple manner, which lends itself under certain conditions to furnishing the basis for a comparison of the reliability of any two tests designed for use in a common situation. An examination of one of the more recent books dealing with the general problem of measurement at the secondary-school level[1] reveals that while the author has assembled in a very careful fashion the available data on the reliability of intelligence tests designed for use with high-school students, he has been forced in his discussion of some widely used tests to resort to general statements instead of specific facts resulting from appropriate statistical analysis.[2] Even in those in-

[1] Odell, C. W. *Educational Measurement in the High School.* New York: Century, 1930. 641 pp.

[2] For example, in a discussion of the Pressey Classification and Verification Tests, Odell has written, "Data on reliability have not been reported, but there appears to be no reason why this should not be almost as reliable as the Miller." (p. 409.)

stances where reliability for a test has been determined by the person by whom it was standardized, or by some later worker, the available results seldom permit a direct comparison of the test with others designed for use in similar situations.

The present study was undertaken to provide comparable reliability figures on five group intelligence tests widely used at the secondary level. For certain of these tests reliability has been determined by the author, or by other persons. Apparently, however, there has been published no direct comparison of the reliability of any two of these tests based upon the results of their administration to a common group of subjects.

The data upon which this study is based were obtained from tests administered to students entering the University High School, University of Minnesota. The tests involved were Army Alpha 8, Pressey Senior Classification, Haggerty Delta 2, Terman Group Test of Mental Ability, Form A, and Miller Mental Ability Test, Form A. All five of these examinations were administered to any given individual upon the same day. In general, groups of twenty-five to seventy-five cases were tested at a sitting. Directions furnished by the authors of the respective tests were followed as closely as possible in administering, scoring and computing intelligence quotients, except in the case of the Pressey Test, where the author's norms were found to be inadequate for superior children. In this instance extrapolation from the author's norms was resorted to as a means of deriving intelligence quotients in those cases scoring above the norms which the author has provided. All children entering the University High School have for some years been given these five intelligence tests. The majority of the cases were tested in April prior to the time of their entrance in September; the remainder were tested in September just before fall-term classes began. Approximately equal numbers of children enter each of the three junior high school grades every year, while a much smaller number is admitted to grade ten.

The data employed in this study have been treated in two groups. Among the test papers of more than one thousand

cases there were represented 192 individuals who were at the time of testing between eleven years, five and one-half months and twelve years, five and one-half months. These cases, which are hereafter referred to as the twelve-year-old group, were tested at a mean age of 143.45 months. There were also 209 sets of tests for individuals who were when tested more than thirteen years, five and one-half months but less than fourteen years, five and one-half months of age. These cases, to be known as the fourteen-year-old group, were tested at a mean age of 166.74 months.

The level of ability represented in the student body of the University High School is considerably above that of typical high-school students. The selective factors affecting the twelve-year-old cases are, of course, of such a nature that distinctly superior ability characterizes this group. The fourteen-year-old group, on the other hand, appears to be of approximately the same level as the student body from which it has been drawn. The variability of each group is restricted beyond that existing in the total population by the selective factors that determine the entrance of students to the University High School. These selective factors are such that the school draws mainly from the upper levels of the occupational scale. On the other hand, the influence of age is in the direction of increasing variability in the raw scores upon which the computations are based. The standard deviations reported in Table I fail to take into account this latter factor, since they are based

TABLE I

Means and Standard Deviations of Intelligence Quotients from Each of Five Tests

TEST	TWELVE-YEAR GROUP (N = 192)		FOURTEEN-YEAR-GROUP (N = 209)	
	Mean	S.D.	Mean	S.D.
Army A	125.31	13.34	118.87	11.44
Pressey	120.71	14.82	113.20	12.69
Haggerty	129.43	18.16	122.14	14.92
Terman	123.75	11.98	114.42	9.66
Miller	133.28	24.82	129.72	19.29

upon distributions of intelligence quotients rather than raw scores.

The above observed standard deviations indicate that in terms of intelligence quotient each group is somewhat less variable than a sample typical of the general population. The fourteen-year-old group shows less variability than does the twelve-year-old group, which includes numerous cases that have been accelerated considerably in their school program.

A comparison of the goodness of fit of each distribution for both age levels is presented in Table II. With the possible exceptions of the distributions for the Pressey and the Terman Tests for the twelve-year-old group, there occur no discrepancies between the observed and the theoretical frequencies that may not be charged to random sampling. However, the P values for the Pressey and Terman Tests for the twelve-year-old group are .0001 and .05 respectively, indicating that causes other than random sampling have affected these distributions.

TABLE II

Goodness of Fit of IQ Distributions from Five Group Tests

TEST	X^2	P IN 100	N'
	Twelve-year-old-Group		
Pressey	52.115	.0001	18
Terman	22.412	.05	14
Army A.	8.924	.92	17
Miller	9.140	.76	14
Haggerty	13.655	.40	14
	Fourteen-year-old Group		
Pressey	19.178	.16	15
Terman	12.616	.70	17
Army A.	19.319	.44	20
Miller	13.372	.64	17
Haggerty	21.001	.23	18

The split-half method has been applied in the conventional manner to determine the reliability of each test. As a first step all of the papers included within the chosen age limits were rescored for the odd and the even numbered items. The correlation between the two sets of half-scores was then computed

for each test, and the Spearman-Brown formula applied to estimate the reliability of the full-length tests. Table III

TABLE III
Reliability of Each of Five Tests

TEST	MEAN SCORE		STANDARD DEVIATION		r_{kk}*
	Odd	Even	Odd	Even	
	Twelve-year Group (N = 192)				
Army A.	46.85	45.40	12.50	13.02	.941
Pressey	19.73	20.73	7.35	7.74	.892
Haggerty	59.94	57.87	9.83	9.98	.862
Terman	57.40	51.50	15.73	15.71	.933
Miller	27.89	28.25	8.84	9.04	.968
	Fourteen-year Group (N = 209)				
Army A.	56.37	54.38	12.16	12.40	.942
Pressey	24.82	26.13	6.98	6.76	.844
Haggerty	66.09	64.58	7.88	7.95	.817
Terman	68.36	64.03	15.00	14.42	.926
Miller	34.10	34.56	7.98	8.46	.937

* Reliability for total test estimated from correlation of odd-even scores by Spearman-Brown formula.

gives the results of this treatment, and permits a direct comparison of the reliability of the tests among groups such as the ones upon which these figures were obtained.

IQ CHANGES IN FOSTER HOME CHILDREN*

EMMETT L. SCHOTT, PSYCHOLOGIST
Henry Ford Hospital

THE purpose of this study was to determine the effect of foster home placement on the IQ of children so placed. Two extensive investigations, dealing with the influence of nature and nurture on the intelligence of foster children, have been reported in the Twenty-Seventh Yearbook of the National Society for the Study of Education. The study made at Chicago (3) emphasized the rôle of the foster home as a determining factor in the test scores of such children. At Stanford University (1) a somewhat different approach to the same problem yielded results emphasizing heredity as a force beside which other influences appeared "dwarfed in comparison."

For our own investigation, 200 children, 100 boys and 100 girls in care of the Michigan Children's Aid Society, were taken in order of their examination in the Division of Neuropsychiatry of the Henry Ford Hospital during a five year period from 1929 to 1934. Most of these children were referred to the placement agency through the Probate Courts of the rural counties of Michigan, and this fact indicates conditions of extreme neglect and dependency in the background of these subjects. Thus, the group as a whole is selected from the lower levels of the socio-economic scale. In chronological age, these children ranged from eighteen months to seventeen

* This paper, in preliminary form, was presented before the Psychology Section of the Michigan Academy of Science, Arts, and Letters in Ann Arbor, Mich., on March 8, 1935. The study was made possible through the liberal cooperation of the Michigan Children's Aid Society staff with headquarters in Detroit. The writer is also indebted to them, to Dr. Thos. J. Heldt, Physician-In-Charge, Division of Neuropsychiatry, Henry Ford Hospital, Detroit, and to Jack H. Schott for suggestions regarding revision of the manuscript.

years, with a median age of 5.6 years at the time of their first test. The Stanford-Binet scale was used on children over three years of age. For younger children, the Kuhlmann extension, the Merrill-Palmer pre-school scale, or the Gesell schedules of development were used. In spite of the selective factor mentioned above, a distribution of the IQ's of these foster children shows the group to be a representative sample of the population, when general intelligence is considered, with nearly normal distribution of scores and with a median falling very close to the IQ norm of 100.

TABLE I
IQ Distribution, First Test

IQ RANGE	CLASSIFICATION	GIRLS	BOYS	TOTAL
50 –69	Defective	4	5	9
70– 89	Retarded	22	23	45
90–109	Average	55	57	112
110–129	Superior	17	14	31
130–149	Exceptional	2	1	3
	Totals	100	100	200
	Median	98.7	97.7	98.2
	Q-1	89.1	87.4	88.2
	Q-3	107.8	106.5	107.1

Of our 200 subjects, seventy-four were retested *by the same examiner,* after placement in foster homes.[1] This is one decided advantage over previous investigations and is in accordance with suggestions made by other authors (4) for the improvement of such studies. The interval between our first and second tests ranged from one month to four and one-half years. The median time between tests was thirteen months. These seventy-four retested cases ranged in age from eighteen months to twelve years, with a median age of three years one month at the time of the first test. Their scores on the two tests were distributed as shown in Table II.

The median scores of these retested subjects are below those of the original group of 200. This probably is partially due to a selective factor operating in the agency referring cases

[1] The term "foster home" includes both boarding and adoptive homes.

TABLE II

IQ Distribution of Retested Group

IQ RANGE	GIRLS		BOYS		TOTAL	
	1st	*2nd*	*1st*	*2nd*	*1st*	*2nd*
70– 89	13	8	18	8	31	16
90–109	16	21	18	24	34	45
110–129	5	4	3	7	8	11
130–149	1	2	0	0	1	2
Totals	35	35	39	39	74	74
Median	95.6	99.0	91.7	99.6	93.5	99.3
Q-1	83.5	90.7	80.8	91.2	81.9	91.1
Q-3	106.6	105.8	102.5	107.7	104.4	107.6

for retest. Children who are failing to adjust in boarding homes are often referred for further examination before permanent placement is attempted, whereas those who are getting along well are less likely to be returned for additional psychological study.

It will be noted that the median IQ of the group increased 5.8 points between the first test and the second test. This difference is not statistically significant when evaluated by the critical ratio technique. The ratio of the actual difference to the standard deviation of that difference is slightly less than .28. However, the slight gain is beyond that expected from "practice effect," according to Cattell (2), even when the two tests are given within an interval of four months. The same investigator reported that very little if any practice effect is carried beyond a six month interval. In the measurement of foster children, the retest after placement probably is a better measure of the child's "best performance" than is the first test at the time of acceptance by the agency. At the time of the first test the child is at a disadvantage, emotionally. He has just been turned over to the agency, usually following a traumatic experience such as loss of a parent, the breaking up of a home, etc. He is in the hospital or temporary boarding home, pending a more permanent disposition of his case and a psychometric examination is requested as an aid in determining suitability for adoption or other placement. His future,

at the time of this first test, is more uncertain than ever before and if he is old enough to realize this, it must have a disquieting influence. A year after placement, when he has regained some sense of security in a foster home, he may more nearly approximate his best performance on a psychometric test. The slight gain of our group as a whole may be due partially to this factor of improved emotional state, but we have no reliable way of measuring such emotional improvement of children in the age range of this group. Neither do we have an accurate rating of the homes in which these children were placed, although the agency considers their foster homes "average or better" with respect to the general population. At the time of acceptance the agency begins a sustained effort to improve the child's physical condition and better health may also be conducive to better test performance.

The coefficient of correlation between the first and second tests was $.82 \pm .026$. Individual changes in IQ scores were distributed as shown in Table III.

TABLE III

Distribution of IQ Changes (1st Test to 2nd Test)

AMOUNT OF CHANGE	NUMBER OF CASES	
Gained 26 to 30 points	2	
" 21 to 25 "	2	
" 16 to 20 "	2	
" 11 to 15 "	2	
" 6 to 10 "	14	
" 1 to 5 "	29	Change of 1 to 5 points 49 cases
No change	8	
Lost 1 to 5 points	12	
Lost 6 to 10 "	3	
Total	74	

Average change regardless of sign 5.4 points.

The retest scores showed more gains than losses and the gains reached a higher range than the losses, but two-thirds of the total number of retest cases changed no more than five points either way. The average change of 5.4 points does not appear to be very significant when compared with the published reports

of variability on retests when no environmental changes aiming at improvement have been introduced into the experiments.

TABLE IV

Comparison with IQ Variation on Retest as Reported in Other Studies

AUTHOR	AVERAGE CHANGE REPORTED
Terman (5)	4.5 points
Rugg and Colloton (5)	4.7 "
Garrison (5)	5.4 "
Rugg, L. S. (5)	3.1 "
Schott, E. L. (7)	
Adult neuropsychiatric patients	7.6 "
Problem children	4.7 "
Foster home children (present study)	5.4 "

Incidentally, nineteen of our seventy-four retested children were tested a third time. A third test was sometimes requested when the results of the second test seemed doubtful or unsatisfactory or when the rating was just below the agency's standard for adoptability or when the child continued to present a problem in adjustment. This group of nineteen had a median IQ of 87 on the first test, 91 on the second test, and 91 on the third test. Thus the slight group improvement between the first and second tests did not continue beyond that, for the third test showed no additional gain. The lack of change on the third test also confirms the validity of the second test as a better measure of the child's true rating than the first test affords, and indicates the advisability of routine retesting of foster children after placement, preferably at an interval of a year or so. One of the greatest values of the foster home lies in enabling the individual child to approximate his own maximum expression of native ability. According to Rogers (6), "a child's behavior may be transformed and his personality changed from a disordered to a cheerful one," although "little change in actual mental capacity can be expected." Our data indicate little hope of raising the IQ appreciably by foster home placement.

SUMMARY

Of 200 children with normal IQ distribution at the time of acceptance by a child placing agency, seventy-four were retested by the same examiner, after placement in foster homes. Median interval between tests was thirteen months. The median IQ of the retested group increased 5.8 points after placement, but two-thirds of the individual scores changed no more than five points either way. The gain after placement is not statistically significant; and the average change is about the same as variabilities reported from investigations unconcerned with environmental improvements affecting the retest. Nineteen cases tested a third time showed no change in the median between second and third tests. Such data indicate little hope of raising the IQ appreciably by foster home placement. However, the variability of mental ratings and the occasional gross change of individuals on retest indicate the advisability of routine retesting of foster children six months or a year after placement.

REFERENCES

1. Burks, B. S., The relative influence of nature and nurture upon mental development; a comparative study of foster parent-foster child resemblance and true parent-true child resemblance. Nature and Nurture; Twenty-Seventh Yearbook of the Nat'l. Soc. for the Study of Ed., Part I, pp. 219–316. Bloomington, Ill. Public School Pub. Co., 1928.
2. Cattell, P., Constant changes in the Stanford-Binet IQ, Jr. Ed. Psychol. 22: 544–550, Oct., 1931.
3. Freeman, F. N., Holzinger, K. J., and Mitchell, B. C., The influence of environment on the intelligence, school achievement, and conduct of foster children. Nature and Nurture; Twenty-Seventh Yearbook of the Nat'l. Soc. for the Study of Ed., Part I, pp. 103–217. Bloomington, Ill. Public School Pub. Co., 1928.
4. Rogers, A. L., Durling, D., and McBride, K., The effect on the intelligence quotient of change from a poor to a good environment. Nature and Nurture; Twenty-Seventh Yearbook of the Nat'l. Soc. for the Study of Ed., Part I, pp. 323–331. Bloomington, Ill. Public School Pub. Co., 1928.
5. Rogers, A. L. et al. (Reference 4) p. 329. Table VII.
6. Rogers, C. R., A good foster home: its achievements and limitations. Mental Hygiene, 17: 21–40; Jan., 1933.
7. Schott, E. L., Variability of mental ratings in retests of neuropsychiatric cases. Amer. J. Psychiat., 10: 213–227, Sept., 1930.

THE INFLUENCE OF AMOUNT LIMITS AND TIME LIMITS UPON THE RATE OF WORK

WILLIAM C. F. KRUEGER
Wayne University, Detroit, Michigan

THIS experiment dealt with two problems: (1) Does knowledge of the amount of work to be done influence the speed of work? What is the rate of work, when it is known that a relatively large amount of work is to be done, in comparison with the rate of work, when it is known that a relatively small amount of work is to be accomplished? (2) What is the speed of work, when it is known that a definite number of units of work is to be completed, in comparison with the rate of work, when it is known that work is to be done for a comparable length of time?

Three groups of subjects were used. These thirty-six subjects were undergraduate students of sophomore and junior rating, who had completed at least the introductory course in psychology and who were relatively untrained in psychological experimentation. The subjects formed a rather homogeneous group in terms of their respective Otis scores (Otis S. A. Test, Higher Examination) and by their respective rates of work as shown by the standard deviations of the calculated average performances.

The work consisted of adding series of one-place numbers for a determined period of time or until a definite number of such problems had been solved. Each problem consisted of two one-place numbers, the sum of which was ten or more. Care was taken that the same problem did not recur twice in immediate succession or close proximity. The subject wrote his answer underneath each problem. Each practice period

was about fifty minutes long. During this practice period a signal was given each fifteen seconds (a tap on the table) and when this signal was given the subject checked the problem he had just completed. Thus amount of work accomplished within a fifteen-second period was indicated by the two successive check marks. There were six practice periods, one practice period in each of six successive weeks. During these weekly practice periods the subjects worked together and drilled to achieve more speed in adding. The work was checked for errors to improve accuracy. In order to get additional training in adding each subject was required to complete a definite amount of the same type of work at home. All work was completed as assigned and within the allotted practice period. Even before all of the practice sessions were completed, gains in efficiency were rather small, and hardly any errors were made. Every subject was kept informed about his progress and achievement.

The same activity was continued during the experiment proper. A carefully arranged counterbalanced order was followed to compensate for the influence of additional practice effect and the usual factors present in an experiment.

Before each practice and experimental work period, the subjects were requested to work at maximum speed at all times. It was also explained that accuracy should not be sacrificed for the sake of speed, nor should too much time be spent to secure 100 per cent accuracy. These requests were repeated occasionally during each period to stimulate and remind the group to maintain a high level of efficiency. Before each of the work periods of the experiment proper the subjects were informed about the length of the work period, either in terms of the number of problems to be solved, or in terms of the temporal length of the particular period.

The six working conditions were: (a) a 5-minute period, (b) a 25-minute period, (c) a 50-minute period, (d) 500 problems, (e) 2,000 problems, and (f) 4,000 problems.

To facilitate the desired comparisons, each work period was divided into successive one-minute periods. The number of

units of work completed or problems solved, for each minute period, was computed. The final results were averaged. No subject who worked under conditions (d), (e), and (f) succeeded in completing his task within the comparable time allotted. For purposes of comparison with conditions (a), (b), and (c) only the completed and corresponding 5-minute, 25-minute, and 50-minute periods for conditions (d), (e), and (f) were used.

Table I contains the averages for the six work conditions with the corresponding S.D.'s and P.E.'s. Since so very few errors were made, these were not included in the treatment of the data. The averages seemed to be fairly reliable as indicated by the small P.E.'s.

The data suggested rather clearly that knowledge of the amount of work to be done influenced definitely the rate of work as far as the initial five-minute periods were concerned. When the amount of work was smaller, more work was done. The differences between the corresponding averages for conditions (a) and (b), also for conditions (a) and (c) were small but of statistical significance throughout the entire five-minute period. The same was true for the differences between corresponding means for conditions (d) and (e), and for conditions (d) and (f). In comparing the averages for the one-minute periods after the initial five minutes, it was found that the smaller amount assigned produced a higher rate of work. While the differences between the means were not as large as in the initial period, for the most part a high degree of statistical significance was maintained.

The answer to the second problem was found by comparing the results for conditions (a), (b), and (c) with the corresponding averages obtained for conditions (d), (e), and (f). It became evident that usually the corresponding averages for conditions (d), (e), and (f) were lower, but the differences between the means were small. While no stress could be placed upon any statistical significance of the difference, it should be noted that conditions (a), (b), and (c) quite consistently yielded higher averages.

TABLE I

Average Scores for Six Conditions of Work

	CONDITION	1ST. MIN.	2ND. MIN.	3RD. MIN.	4TH MIN.	5TH MIN.	6TH MIN.	7TH MIN.	8TH MIN.	9TH MIN.	10TH MIN.	11TH MIN.	12TH MIN.	13TH MIN.
(a)	Av.	84.5	79.6	78.8	78.0	79.0								
	S.D.	6.3	6.9	6.8	6.5	6.8								
	P.E.	.69	.76	.75	.71	.75								
(b)	Av.	82.6	75.6	75.4	73.9	74.2	73.3	73.0	72.9	72.8	72.5	72.3	72.1	72.2
	S.D.	6.8	7.5	7.0	6.8	6.3	5.5	5.7	5.5	5.0	6.5	6.1	6.1	6.2
	P.E.	.75	.82	.77	.75	.69	.60	.63	.60	.55	.71	.67	.67	.68
(c)	Av.	79.5	72.1	71.9	71.0	71.3	70.8	70.0	70.2	69.7	69.5	69.2	69.0	69.1
	S.D.	7.0	6.7	6.3	6.8	7.0	6.8	6.4	6.3	5.8	6.4	5.9	6.1	6.2
	P.E.	.77	.74	.69	.75	.77	.75	.71	.69	.64	.70	.65	.67	.68
(d)	Av.	82.0	79.0	78.6	78.1	78.1								
	S.D.	7.4	6.7	6.2	6.5	7.0								
	P.E.	.81	.74	.68	.71	.77								
(e)	Av.	77.8	74.5	73.8	73.6	73.2	72.8	72.7	72.1	72.2	72.2	71.8	71.2	71.5
	S.D.	6.5	6.1	6.9	7.0	5.9	6.8	6.3	6.7	6.1	6.4	6.0	5.7	5.9
	P.E.	.71	.67	.76	.77	.65	.75	.69	.74	.67	.70	.66	.63	.65
(f)	Av.	75.4	71.8	71.2	71.2	71.0	70.4	71.0	70.2	69.2	69.2	69.2	69.0	68.8
	S.D.	6.8	7.3	7.5	7.1	6.9	6.3	6.4	6.8	6.7	6.9	6.5	6.3	5.9
	P.E.	.75	.81	.82	.78	.76	.69	.70	75	.74	.76	.71	.69	.65

TABLE I—(*Continued*)

	CONDITION	14TH MIN.	15TH MIN.	16TH MIN.	17TH MIN.	18TH MIN.	19TH MIN.	20TH MIN.	21ST MIN.	22ND MIN.	23RD MIN.	24TH MIN.	25TH MIN.	26TH MIN.
	Av.	72.4	71.9	72.0	71.6	71.4	71.7	70.9	69.8	70.2	69.5	69.3	69.3	
(b)	S.D.	6.2	6.3	6.3	6.3	5.1	5.3	6.1	4.6	5.1	5.4	6.4	5.7	
	P.E.	.68	.69	.69	.69	.56	.58	.67	.51	.56	.59	.70	.63	
	Av.	68.9	69.0	68.4	68.6	68.6	68.6	68.5	68.4	68.2	67.8	68.0	67.2	67.3
(c)	S.D.	6.4	6.7	6.7	6.2	5.8	5.9	6.1	5.4	5.8	5.7	5.9	5.8	6.0
	P.E.	.70	.74	.74	.68	.64	.65	.67	.59	.64	.63	.65	.64	.66
	Av.	71.5	71.1	71.7	70.9	70.7	70.2	70.0	69.9	69.3	69.4	69.2	69.8	
(e)	S.D.	6.0	6.7	6.1	6.8	6.1	6.0	5.7	5.4	5.9	5.9	6.5	6.8	
	P.E.	.66	.74	.67	.75	.67	.66	.63	.59	.64	.64	.71	.75	
	Av.	68.4	68.8	68.6	68.4	68.2	68.0	67.6	67.8	67.4	67.2	67.0	67.3	66.7
(f)	S.D.	6.2	6.0	7.0	6.5	6.2	5.6	6.0	6.4	5.9	5.5	5.8	6.1	6.0
	P.E.	.68	.66	.77	.71	.68	.62	.66	.70	.65	.60	.64	.67	.66

	CONDITION	27TH MIN.	28TH MIN.	29TH MIN.	30TH MIN.	31ST MIN.	32ND MIN.	33RD MIN.	34TH MIN.	35TH MIN.	36TH MIN.	37TH MIN.	38TH MIN.	39TH MIN.
	Av.	67.0	66.1	66.7	66.3	65.2	65.8	65.6	65.1	64.8	64.9	65.0	64.7	64.3
(e)	S.D.	6.3	6.1	6.0	6.0	5.2	5.5	6.1	5.4	6.1	5.3	5.8	5.9	6.1
	P.E.	.69	.67	.66	.66	.57	.60	.67	.59	.67	.58	.64	.65	.67
	Av.	66.0	66.0	65.5	65.2	65.9	65.0	65.4	65.3	65.0	64.3	64.3	64.3	63.7
(f)	S.D.	5.9	5.4	5.9	6.0	5.7	6.0	6.1	5.4	6.1	5.3	5.8	5.9	6.1
	P.E.	.65	.59	.65	.66	.63	.66	.67	.59	.67	.58	.64	.65	.67

	CONDITION	40TH MIN.	41ST MIN.	42ND MIN.	43RD MIN.	44TH MIN.	45TH MIN.	46TH MIN.	47TH MIN.	48TH MIN.	49TH MIN.	50TH MIN.
	Av.	64.3	64.1	63.4	62.9	62.7	62.3	62.0	61.3	60.7	60.4	59.3
(e)	S.D.	6.1	6.2	5.8	6.3	6.2	5.9	6.0	6.1	5.3	5.6	6.2
	P.E.	.67	.68	.64	.69	.68	.65	.66	.67	.58	.62	.68
	Av.	63.7	63.2	62.5	62.7	61.8	60.8	60.3	60.5	59.9	58.7	58.7
(f)	S.D.	6.1	5.8	5.4	5.6	5.5	5.2	5.6	5.6	5.9	6.0	5.4
	P.E.	.67	.64	.59	.62	.60	.57	.62	.62	.65	.66	.59

The subjects at first complained that they were soon tired of the work as well as by it. The difficulties were attributed to ocular fatigue and fatigue in the hand used for writing. During the experiment the most serious difficulty was that the hand and arm became painfully tired toward the end of the long work periods. The subjects reported that they preferred the shorter periods and almost dreaded those long and monotonous fifty-minute periods. During the earlier practice periods, the subjects had to stop work to shake and stretch the tired hand a little before continuing to work. This interruption never occurred during the regular experimental work periods. The subjects also frequently complained about ocular fatigue during the early practice sessions. The numbers seemed blurred and it was difficult to observe them in the regular positions. While the hand still became very tired in the long work periods, the complaints about ocular fatigue were not made during the experiment proper.

CONCLUSIONS

The data obtained under the conditions of the experiment seemed to justify the following conclusions:

(1) The rate of work was higher for the relatively smaller work period. This held true when the amount of work was defined by the number of problems to be completed, and when the task was stated in terms of the temporal length of the work period.

(2) The differences between the rates of work were smaller when the relative proportions of the work to be done were smaller.

(3) The differences between the rates were larger during the earlier part of work period.

(4) The rate of work was rather consistently higher when the task was stated in terms of the number of units of work to be completed than when the task was determined by the temporal length of the work period.

THE EFFECTS OF TRAINING UPON THE BELIEF IN CERTAIN POPULAR MISCONCEPTIONS

RICHARD S. HARTER
Temple University

PREVIOUS STUDIES

SEVERAL investigators have found that belief in certain popular misconceptions or superstitions is rather common among college students. Conklin (1) found in a group of over 500 college students that 90 per cent of the women and 73 per cent of the men could recall belief in or practice of superstitions at some time or other. The women were inclined to have superstitions concerning the home, social relations, wishes, love, marriage, and death; whereas, most of the superstitions of the men concerned sports and business activities.

The prevalence of belief in certain misconceptions among college students was investigated by Nixon (2), who submitted a questionnaire of 30 popularly endorsed statements (none of which was supported by scientific fact) to 140 women and 219 men students in elementary psychology at Columbia and New York Universities. Nixon's results indicated a greater credulity among the women, as evidenced by a tendency for them to accept more of the statements as true than the men. The order of merit for belief in the 30 statements was similar for the men and women, as shown by the rank order coefficient of 0.90.

Garrett and Fisher (3) added ten more false popular statements to Nixon's questionnaire and presented it to 140 senior girls in the Washington Irving and to 115 senior boys in the DeWitt Clinton High Schools in New York City. Their re-

sults, also indicating that the girls were slightly more credulous than the boys, show a small but reliable sex difference. There was very little difference in the order of merit of belief in the 40 statements for the boys and girls, as indicated by the rank order coefficient of 0.92. Intelligence records were available for 100 boys and 100 girls, so correlations were obtained between intelligence and the misconception scores. Negligible correlations were reported, r being -0.02 for the boys and -0.11 for the girls.

Gilliland (4), in his study of the effects of a course in general psychology upon common superstitions and prejudices, also used a modified form of the Nixon questionnaire. Two groups of students in general psychology were selected: 103 students from the College of Liberal Arts, and 46 students in the evening School of Commerce at Northwestern University. The questionnaire was given to each group at the first meeting of the class and again at the last meeting. In the first test, the better students were prone to be less credulous than the poorer students. The results of the second test showed that training was more advantageous to the better students, there being a greater reduction in credulity for this group.

PURPOSE OF THIS STUDY

The purpose of this investigation is to study the effects of psychological training upon belief in certain misconceptions and to determine, if possible, the relative importance of some of the factors responsible for the acceptance of these beliefs.

METHOD OF PROCEDURE

In this study the Garrett and Fisher questionnaire (a modified form of Nixon's questionnaire) was given at the beginning and again at the close of the semester to 94 women and 154 men students in elementary psychology at Temple University. At the first presentation of the questionnaire (first class session), the students were told to answer the statements to the best of their ability. There was no specific time limit set. At

the second presentation of the questionnaire (last class session), the students were asked to state their reasons for accepting the statements as true or false.

QUESTIONNAIRE USED IN THIS STUDY

Directions: Below are a number of statements. Some of them you may consider true, some false. If you think a statement is substantially true, draw a circle around the T.; if you think the statement false, draw a circle around the F. Where it seems debatable, mark as seems to you nearest correct.

1. T.F. The number of man's senses is five.
2. T.F. A child comes into the world with an instinctive knowledge of good and evil.
3. T.F. Certain lines in a person's hand foretell his future.
4. T.F. If you stare at a person's back, you can make him turn around. This is a form of telepathy.
5. T.F. It is unlucky to have anything to do with the number 13.
6. T.F. A man's character can be read by noting the size and location of certain special developments on his head.
7. T.F. People with greenish eyes are not as trustworthy as people with blue or black eyes.
8. T.F. An expectant mother, by fixing her mind on a subject, can influence the character of her unborn child.
9. T.F. Women are inferior to men in intelligence.
10. T.F. People born under the influence of certain stars show this influence in their characters.
11. T.F. Intelligence can be increased by proper training.
12. T.F. Long slender hands indicate an artistic nature.
13. T.F. Beginning an undertaking on Friday is almost sure to bring bad luck.
14. T.F. If a man had faith enough, he could heal a broken limb.
15. T.F. Many eminent men have been feeble-minded in childhood.
16. T.F. Some animals are as intelligent as the average man.
17. T.F. No defect of body or mind can hold us back if we have will power enough.
18. T.F. Adults sometimes become feeble-minded from overstudy.

19. T.F. All men are created equal in capacity for achievement.
20. T.F. The marriage of cousins is certain to result in children of inferior intelligence.
21. T.F. Especially intelligent children are apt to be weak and retarded physically.
22. T.F. The study of mathematics is valuable because it gives one a logical mind.
23. T.F. You can estimate an individual's intelligence pretty closely by looking at his face.
24. T.F. A square jaw is a sign of will power.
25. T.F. A high forehead indicates superior intelligence.
26. T.F. A person who does not look you in the eye is likely to be dishonest.
27. T.F. Women are by nature purer and nobler than men.
28. T.F. Man is superior to the animals because his conduct is guided by reason.
29. T.F. Any physical or mental disease can be contracted by thinking too much about it.
30. T.F. Blondes are less trustworthy than brunettes.
31. T.F. Cold hands are the sign of a warm heart.
32. T.F. A person who holds his thumbs in his hands is a coward.
33. T.F. If your ears burn, it is a sign that some one is talking about you.
34. T.F. Conscience is an infallible guide to conduct.
35. T.F. Silent men are generally deep thinkers.
36. T.F. Fat people are always good natured.
37. T.F. Women possess a power of intuition absent in men.
38. T.F. Chess and checker playing develops one's power of concentration.
39. T.F. The slow learner retains better what he gets than the quick learner.
40. T.F. Brains and beauty rarely go together.

TREATMENT OF THE RESULTS

The distributions in Table I show how many of the statements were accepted as being true by the men and women for the two presentations.

An analysis of this table indicates that at the first presentation (before training in psychology), the men marked 9.4

TABLE I

Number of the 40 Statements in the Questionnaire Marked True for the First and Second Presentations

NUMBER OF STATEMENTS MARKED TRUE	FIRST PRESENTATION		SECOND PRESENTATION	
	Men	Women	Men	Women
29	1			
28				
27	2			
26				
25				
24	2			
23		1		
22	3	2		
21	1	4		
20	5	2	1	
19	4			
18	2	3	1	
17	4	2		
16	2	2		1
15	8	3		
14	3	3		2
13	9	5	2	
12	7	5	2	1
11	11	4	8	4
10	6	8	1	2
9	6	5	5	3
8	10	3	5	3
7	8	6	9	4
6	7	4	6	5
5	15	9	3	1
4	9	3	7	3
3	7	3	8	8
2	4	5	22	7
1	5	6	21	16
0	13	6	53	34
Total	154	94	154	94
Average (Mean)	9.4	9.2	3.2	3.2
S.D.	6.5	6.2	4.0	3.6
S.D. Average	.52	.63	.25	.32
S.D. Difference	.81		.4	
Experimental Coefficient (McCall)	.08		.11	

(mean) of the statements true; the women, 9.2 (mean) statements true. At the second presentation (after training in psychology), the means were 3.2 for the women and 3.2 for the

men. It is interesting to compare these results for the first presentation with the results obtained by Nixon and by Garrett and Fisher. In Nixon's study the men marked 35.1 per cent of the statements true; the women, 41.1 per cent. In the study of Garrett and Fisher the boys marked 44 per cent of the statements true; the girls, 49.7 per cent. In our study the men marked 23.5 per cent of the statements true, and the women 23 per cent. Our findings, however, do not reveal any significant sex differences in credulity. The difference in credulity is so negligible between the 94 women and the 154 men, that the experimental coefficient (McCall's device for showing the reliability of a difference) averages only 0.08 for the first presentation and 0.11 for the second. An experimental coefficient of 1.0 would be needed to give a reliable sex difference. Furthermore, it is of interest to note (Table III) that regardless of certain variations, the rank order of belief in the 40 statements corresponds closely for the sexes, r being 0.91 for both the first and second presentations. The standard deviations (Table I) show that the degree of variability is about the same for each sex.

Intelligence scores and class grades (derived from objective test scores) were obtained for 150 students. Correlations were computed between the indicated variables represented in Table II.

TABLE II

Correlation between Intelligence, Class Grades, and Misconception Scores

	1st R	2nd R
Between intelligence and misconception scores	−.23	−.22
Between class grades and misconception scores	−.23	−.20

Certainly it is apparent from these low negative correlations that there are very slight relationships existing between the degree of acceptance of these misconceptions and the factors of intelligence and scholarship.

The effect of training upon belief in these misconceptions was another important factor considered in the analysis of this investigation. It should be kept in mind, however, that the training received by the students in elementary psychology did not include specific instruction with regard to such statements as those numbered 11, 15, 18, 20, 21, 22, 28, 34, 38, and 39. Consequently, the data presented in Table III which shows the percentage of each group marking each statement true, manifest a greater percentage of acceptance for these statements, due to the fact that the students were obliged to rely upon previously acquired associations or miscellaneous sources of information in answering them. Despite this disadvantage, a comparison of the results between the first and second presentations indicate that there was considerable reduction in credulity, especially for those statements for which correct scientific information had been provided.

For example, it is interesting to note (Table III) that for the first presentation, 80 per cent of the men and 72 per cent of the women believed that "The number of man's senses is five." This was the most popularly accepted statement for the men, and second in popularity among the women. For the second presentation, however, only 10 per cent of the men and 15 per cent of the women accepted this statement as true. The popularity or order of belief for the statement was reduced to the 9.5 rank by the men and to the 7.0 rank by the women.

While there was an acceptance of approximately 23 per cent of the statements by the entire group of subjects at the first presentation, it was significant that there was a reduction to an acceptance of only 8 per cent of the statements at the second presentation. Nevertheless, certain of these beliefs seem to persist to a considerable extent despite education and scientific training.

This fact, and also the negligibility of the factors of intelligence and quality of scholarship in the determination of these beliefs, shows that other influences are involved. Some idea of the significance of these influences was evolved from a care-

TABLE III

Rank Order of Relief Among the Students for the 40 Statements for the First and Second Presentations

STATEMENT NO.	IDEA INVOLVED	1ST PRESENTATION				2ND PRESENTATION			
		Per cent marking true		Order of Prevalence		Per cent marking true		Order of Prevalence	
		M.	F.	M.	F.	M.	F.	M.	F.
1	Man's senses are five	80	72	1	2	10	15	9.5	7
28	Man superior to animals in reason	74	93	2	1	30	45	4	1
38	Chess and checker playing develop power of concentration	72	53	3	4	40	34	1	2
11	Intelligence increased by training	70	61	4	3	36	26	2	3
22	Mathematics gives logical mind	58	51	5	5	31	23	3	4
34	Conscience infallible guide conduct	45	40	6	6	18	19	5.5	5
18	Adults feeble-minded from study	33	29	7	10	15	15	7	7
17	Will power supreme	31	34	8	8	7	11	16	10.5
12	Artistic nature shown by hands	30	36	10	7	8	7	13	16.5
35	Silent men usually deep thinkers	30	19	10	18.5	8	11	13	10.5
39	Slow learner a good retainer	30	19	10	18.5	18	10	5.5	12
4	Telepathic influence of staring	29	33	12	9	3	9	25	13.5
15	Early feeble-mindedness of great men	24	18	13	21.5	10	6	9.5	19
23	Face indicates intelligence of person	23	21	14	14	7.7	4	15	22
16	Some animals intelligent as humans	22	16	16	25	1.3	2	29	28
37	Women possess power of intuition	22	25	16	12	6.5	15	17.5	7
6	Character shown by head shape	22	20	16	16	.6	2	33	28
29	Thinking may cause disease	21	17	18	23.5	9.7	9	11	13.5
27	Women purer by nature than men	19	21	19	14	6.5	7	17.5	16.5

TABLE III—(*Continued*)

STATEMENT NO.	IDEA INVOLVED	1ST PRESENTATION				2ND PRESENTATION			
		Per cent marking true		Order of Prevalence		Per cent marking true		Order of Prevalence	
		M.	F.	M.	F.	M.	F.	M.	F.
21	Bright children physically retarded	18	18	20.5	21.5	11	6	8	19
26	Shifty eye shows dishonesty	18	26	20.5	11	5	6	20	19
40	Brains and beauty rarely go together	16	7	22	30.5	3.2	4	23.5	22
36	Fat people are always good natured	15	8	24.5	29	5	4	20	22
20	Marriage of cousins gives defectives	15	19	24.5	18.5	5	13	20	9
24	Square jaw is sign of will power	15	19	24.5	18.5	2.6	1	26	31
19	Men created equal in capacity	15	21	24.5	14	8	8	13	15
9	Women inferior to men in intelligence	12	3	27	35	3.2	0	23.5	36.5
25	High forehead shows intelligence	11	13	28.5	26	1.3	3	29	25
14	Faith can heal a broken limb	11	6	28.5	33	3.9	1	22	31
2	Instinctive knowledge of good	10	7	31	30.5	1.3	3	29	25
8	Prenatal influence	10	11	31	27	1.3	2	29	28
10	Character influenced by stars	10	17	31	23.5	1.3	3	29	25
3	Lines in hand foretell future	8	10	33	28	0	0	37.5	36.5
30	Blondes less trustworthy	7	0	34	39	.6	0	33	36.5
31	Cold hands indicate a warm heart	4	6	35.5	32	.6	1	33	31
33	Burning ears means someone talking	4	4	35.5	34	0	0	37.5	36.5
7	Green-eyed people untrustworthy	2	1	37	36.5	0	0	37.5	36.5
32	Coward holds thumbs in fists	.6	1	38	36.5	0	0	37.5	36.5
5	Thirteen unlucky	0	0	39.5	39	0	0	37.5	36.5
13	Friday unlucky	0	0	39.5	39	0	0	37.5	36.5

ful study of the reasons given by the students for acceptance of the statements.

An analysis of these reasons indicates that belief in the misconceptions is sustained by the following influences: lack of adequate information, religion, palmistry, astrology, phrenology, physiognomy, etc. Lack of information concerning the scientific meaning of such terms as heredity, intelligence, feeblemindedness, insanity, transfer of training, telepathy, intuition, etc., seems to influence the acceptance of the statements in the majority of cases. It is apparent in many cases that various religious teachings exert an influence in the acceptance of many of these misconceptions. Although the students seem to know the scientific meaning of terms and to have the scientific point of view, they seem to prefer a religious or idealistic interpretation of the problem.

This tendency was rather pronounced and it was perhaps mainly responsible for the low negative correlations obtained between the degree of acceptance of the misconceptions and the factors of intelligence and scholarship. That is, the most intelligent and scholarly students often reacted to the situations according to personal feeling or desire, rather than in an objective sense. Students acknowledged, in many of their reasons, that science said "No" to a problem, but that they were willing to say "Yes," or to accept a statement as true, because of some particular personal experience or feeling.

It is probably true that past training and experience have been most effective in the establishment of specific and more permanent modes of response to many situations, and that these are carried over into the adult period and retained because they possess emotional value for the individual. Then, too, it should be remembered that many students of today are impressed and influenced by the idealistic interpretations of natural phenomena advanced by such well-known scientists as Jeans, Compton, Eddington, and others. Thus, many of the reasons offered by the students seem to imply that they want to believe at times in these and similar compromising solutions.

It might be said in conclusion that most students accept the authority of science in the majority of situations, but in a few personal relations prefer some form of idealistic interpretation. The results of this study seem to justify, in part, a definition of reasoning by Clarence Darrow: "Reasoning is simply a method of comparing and appraising; it is always used to justify what the emotions demand."

REFERENCES

1. CONKLIN, EDMUND S. "Superstitious Beliefs and Practices Among College Students," *American Journal of Psychology*, 1919, 30, 83–102.
2. NIXON, H. K. "Popular Answers to Some Psychological Questions," *American Journal of Psychology*, 1925, 36, 418–423.
3. GARRETT, H. E., AND FISHER, T. R. "The Prevalence of Certain Popular Misconceptions," *Journal of Applied Psychology*, 1926, 10, 411–420.
4. GILLILAND, A. R. "A Study of the Superstitions of College Students," *Journal of Abnormal and Social Psychology*, 1930, 24, 473–479.

NEWS AND NOTES

The Midwestern Psychological Association will hold its twelfth annual meeting at the University of Illinois, Urbana, Illinois, on May 7 and 8, 1937, under the presidency of Dr. Harvey A. Carr. The title of Dr. Carr's address will be "The Search for Certainty."

Dr. J. E. W. Wallin, Delaware Director of Special Education and Mental Hygiene, has been made Secretary of the Delaware Sesquicentennial Commission. He was elected a member of the Board of Directors at the recent annual Convention of the International Council for Exceptional Children and appointed a member of its committee on teacher training. Dr. Wallin will offer graduate courses on mental deficiency and mental hygiene in the first and third terms of the summer school of Duke University.

The World Book Company has just published two interesting tests for use in the junior and senior high school. The first is the Thompson Business Practice Test, by James M. Thompson, of New York University. The primary purpose of the test is for measuring achievement and analyzing students' accomplishment in those general business practice courses which are offered under such names as junior business training, introduction to business, or everyday business. The second is the Kefauver-Hand Guidance Tests and Inventories, composed of six tests and two inventories. The tests relate to the educational, social-civic, health, recreational, and vocational activities of the student. They provide an objective basis for determining how well students are informed of a wide range of opportunities in each of these phases of his life. The two inventories indicate whether specific plans or goals have been formulated by the student, and how accurately he defines his own abilities.

The National Bureau of Casualty and Surety Underwriters, 1 Park Avenue, New York, has just published a 48-page manual entitled "A Teacher's Manual—Designed for Use with 'Man and the Motor Car,'" the latter a 256-page textbook published a year ago in response to urgent requests from many high schools which have now inaugurated automobile driving instruction or traffic safety as part of their regular curriculum. The Teacher's Manual, prepared under the editorial direction of Dr. Herbert J. Stack, Director of the National Bureau's Education Division,

contains 16 units of instruction, each one of which presents a group of text questions, several problems relating to the lesson, a number of student activities, and a broad list of supplementary references.

A summary of the Thirty-first Annual Report of the President and Acting-Treasurer of the Carnegie Foundation for the Advancement of Teaching contains interesting excerpts from articles appearing in the Annual Report. Dr. Walter A. Jessup, the president of the Foundation, contributes an article, "The Exploitation of Youth," in which he says: "Society . . . on the whole, has continuously intervened between the child and his parent, between the child and industry and between the child and government. The social object has been to protect the child from abuse and neglect during his period of weakness. In America this has been clearly recognized through the elaborate development of our public schools."

Dr. Alfred Z. Reed, staff member, discusses the contribution of the medieval university to American higher education. For grants for eighteen special research projects in education during the year ended June 30, 1936, Carnegie Corporation of New York paid to the Division of Educational Enquiry of the Carnegie Foundation $94,500. Since 1905 the Foundation has disbursed more than $30,800,000 in retiring allowances and widows' pensions. Copies of this Report may be had without charge on application to the office of the Foundation, 522 Fifth Avenue, New York City.

Founded three years ago as a clearing house for authoritative information on the problems of the exceptional child, the Child Research Clinic of The Woods Schools of Langhorne, Pa., has annually sponsored conferences in leading eastern cities in addition to Institutes on the Exceptional Child at Langhorne, Pa. On February 27, 1937, a conference was held at Washington, D. C. Among the speakers was included Dr. Winifred Richmond, psychologist at St. Elizabeth's Hospital, Washington, who spoke on "The Exceptional Child and the Family Constellation." On March 20 another conference will be held in New York City. Speakers on the program will include Dr. Goodwin Watson, Columbia University; and Dr. Frank Astor, liaison officer between the National Child Welfare Association and the Bureau of Child Guidance, Board of Education, New York City.

BOOK REVIEWS

WINIFRED V. RICHMOND. *Personality: Its Development and Hygiene.* Farrar & Rinehart, 1937.

Like practically every other book which sets out to be a text on personality, Dr. Richmond's volume lays heavy stress on the pathological side of human behavior. It seems to me that this represents a serious misconception of the psychology of personality. In sheer quantity the normal personalities far outnumber those who are abnormal; and they too have a genesis and a development!

From a psychologist's viewpoint, the book seems unevenly written. At times it seems destined for a popular audience, with few technical terms and the factual material spread very thinly. At other times the material (*e.g.*, on Freudian mechanisms) is presented in very compact form and in terminology which the average student, with only a general psychology course as background, will find difficult reading.

The insertion of an entire chapter on intelligence seems rather unjustified. The intellectual aspects of behavior are generally thought to be contributory rather than essential phases of personality. This chapter could have profited by the assistance of a more "academic" psychologist.

A chapter on conditioned reflexes and habits is a pleasure to find, and it is well done. The importance of feelings and emotions, the development of affective habits, etc., are properly stressed. This feeling of pleasure disappears when the reader searches the remaining chapters for systematic use of the concepts of feeling and emotion, conditioned reflex, habit, etc., as explanatory concepts in the analysis of personality development. I was unable to find any cases in which the learning sequence in the development of neurotic symptoms, for example, was analyzed in any detail. While the intelligent student may correlate the earlier and later chapters, the average student is not likely to make this transfer.

The conclusions on heredity and environment show a better balance than persons with medical bias usually attain. Many persons with biological orientation, such as Dr. Richmond manifests, overestimate the possibilities of heredity, eugenics and sterilization. This book is conservative in its conclusions on that point.

In emphasizing the negative criticisms of Dr. Richmond's volume, I do not wish to give the impression that it has no good points. I think it is fairly representative of much work being done, and the point of view of a considerable number of students of personality. On the other hand, I

protest the majestic wave of the hand with which all measurement approaches, the concept of traits, etc., are cast aside. This attitude is unfortunately characteristic of a large proportion of clinical psychologists, an attitude which seems quite unfair. The location of the book in the curriculum would be on the borderline between general psychology and mental hygiene. It does not fit the latter because of its slender discussion of methods of treatment. I should not assign it to a separate course in personality because it does not contain much material which is not treated in either the general or the mental hygiene course.

ROSS STAGNER,
University of Akron.

DUDLEY D. SHOENFELD, M.D., formerly adjunct-psychiatrist, Mount Sinai Hospital, New York. *Crime and the Criminal.* 411 pages. Covici Friede. New York. $3.50.

This intriguing book is dedicated to the theory that crime is an illness and that psychiatry, therefore, has an important place in the administration of criminal law—not only in the correction of asocial tendencies of the criminal after conviction but also in his detection and apprehension through postulating his motives and "mental make-up." Although the author's approach is fundamentally Freudian, Gestalt psychologists will be interested in his statement that we must not only consider the criminal but also his relationship to the victim and, if we are to arrive at an understanding of the crime, discover what factors in the latter make him vulnerable to attack.

At present the law asks of the psychiatrist only his opinion as to the sanity of the accused by which it means, as in 1843, whether he is able to tell the difference between right and wrong. The advance in psychiatry since that date, despite "the battles of experts" in our courts, is largely ignored; however in the Lindbergh case the author was allowed to demonstrate modern applications of his science. Called into the case within a month of the kidnapping, his theoretical deductions from the study of the ransom notes did much to guide the police, suggesting that the criminal was a lone worker, rather than a member of a gang as was generally believed, that he was a German, living in a certain Bronx district, about forty years old, of mechanical occupation, and that he would always carry as a fetish some of the ransom money—predictions which were fulfilled by the arrest of Hauptmann. Dr. Schoenfeld says that according to the simple legal test Hauptmann was sane but was in his opinion a case of dementia praecox; he pleads that a thoroughly scientific study of his crime would have proved of more value to society than his vindictive execution and regrets that a more profound psychiatric analysis could not have followed the two superficial examinations that were actually permitted. The book gives not only a detailed account of the detailed steps

leading to the arrest of Hauptmann but also a condensed version of the trial (the transcripts fill twelve books) as seen through the eyes of a psychiatrist. Teachers of the psychology of crime will be unable to avoid reference to the Lindbergh case for years to come and in this study they will find their best source book.

LEWIS HERMAN AND MEYER GOLDBERG, Member of the Chicago Bar. *You May Cross Examine.* New York. 1936. The Macmillan Co. Pp. 194. $2.00.

A New York bus driver after helping to the sidewalk fifteen of his gas-stricken passengers ran to summon aid; ambulance doctors found no less than fifty stretched out for medical attention. Such attitudinal behavior being common and perjury more and more prevalent since the weakening of the bonds of the oath, the administration of justice has become increasingly difficult. Emphasis upon cross examination has grown accordingly, and the law concedes the utter unreliability of testimony that has not been so tested.

The authors admit that basic psychological precepts, which must be considered in any analysis of cross examination, underlies the entire structure of court procedure; however, they believe that the courts are 'far in advance of the so-called modern science of psychology.' This is comprised for them by Munsterberg's, *On the Witness Stand,* of 1923, McCarty's, *Psychology for the Lawyer,* of 1929, and Burtt's, *Legal Psychology,* of 1931. Such terminology as memory, attention, and association of ideas still finds employment here with the interjection of a few Freudian terms such as 'wishful thinking.' There is a startling chapter on sex differences in testimony which some graduate student may be interested in proving or disproving. Forensic psychologists may be further surprised at the information that 'the latest wrinkle in psychology is the belief of certain professors that human emotions may be classified and that each emotion is betrayed by various external movements or involuntary gestures of the hands' and that a group of Eastern psychologists has made a list of thirty-five such emotions and their accompanying gestures which a shrewd lawyer may use to confound lying witnesses.

"The courts would be in a bad way indeed," say the authors, "if they were given over to modern psychologists, to be used as playgrounds for their block building and paper cutting tests. A commission of experts would probably be established to give every witness a complete series of tests, so as to determine the efficiency of this perception, power of memory, and his ability to narrate. Probably a whole series of scientific statistics would be worked out to determine if the witness was normal, from a psychological standpoint, so as to make his testimony competent. But the fun would begin if the commission of experts determined in their

wisdom that the witness was not competent, and he produced his own group of experts who were ready to swear that he was perfectly capable of testifying."

"Besides analyzing and pointing out possible sources of error and contradiction in the testimonial process, modern psychologists have evolved various instrumentalities which they hope will prove useful as an adjunct to cross examination," say the authors in referring to the much publicized lie detector, whose use is now confined to investigations before trials. This they admit has secured confessions in some cases, probably because of fear, but a United States Court has ruled that it has not gained such standing and scientific recognition among physiological and psychological authorities as would justify the courts in admitting its evidence. A poll of the American Psychological Association is quoted as having brought forth a few responses,—thirteen positive that the tests were not practical, seven doubtful, and eighteen feeling that the tests were useful with certain qualifications, such as the fact that they should be administered only by experts, which the authors take as another example of contradiction in expert opinion. Scopolamine is mentioned as of doubtful value and even more 'claptrap' are the claims for hypnotism. (Sumner's use of the psychogalvanometer had not been made public at the time the book was written.)

Research possibilities in the court room have, of course, been neglected by psychologists, who will find stimulating reading in this book. In many quoted cases Fallon, Rogers, Leibowitz, and other famous lawyers are portrayed as clever amateur psychologists, as in Leibowitz's dramatic proof in open court that a witness in an arson case who swore that he had smelled kerosene was unable to distinguish between vials of kerosene, gasoline, gin, and water. Psychologists may become sought as expert advisors to cross examination and eventually establish more scientific procedures for the determination of justice. That the present reliance of the courts upon cross examination is dangerous is shown by one case where witnesses put to flight by able lawyers were proved later to be not perjurers but truthful citizens whose confusion had brought about a miscarriage of justice. The lesson of Borchard's, *Convicting the Innocent*, should not be quickly forgotten. "The time is ripe," says Sheldon Glueck in his recent *Crime and Justice*, "for a radical re-examination and overhauling of the complicated and ill arranged structure of criminal justice," and *You May Cross Examine* will serve to remind psychologists of their opportunities in this field.

JAMES HARGAN,
Psychologist, Sing Sing Prison

SKINNER, CHARLES E., *Editor*. *"Educational Psychology."* New York: Prentice-Hall, Inc. 1936. Pp. XXV + 754.

The reviewer has used this book as a text before writing the review. Each of the twenty-five chapters of the book is written by a different writer. The names of most of them are familiar to students in the field. The volume attempts to cover the whole field of educational psychology with the result that it is probably too extensive for use in those institutions which have specialized courses for segments of the subject. To cover all of the ground presented in this volume in a three semester hour course, for example, would be a very large assignment. Although each author was assigned a special subject, it was apparently impossible for each to realize just what other writers were including and there is, therefore, some duplication of material which would undoubtedly not have existed had the whole treatise been written by one person. For the most part the content of the book is straightforward and non-technical so that it can be read easily by those who have little, if any, background in the subject. For those who wish to go into the matters more completely, there are in most of the chapters abundant references to technical studies. On the whole, the reviewer has found the volume satisfactory as a text.

D. A. WORCESTER,
University of Nebraska

NEW BOOKS AND PAMPHLETS

Books and pamphlets for review should be sent to James P. Porter, *Editor*, JOURNAL OF APPLIED PSYCHOLOGY, Ohio University, Athens, Ohio.

Les Aptitudes de L'Homme. C. SPEARMAN. Publications du *Travail Human.* Conservatoire National. Des Arts et Metiers, Paris, 1936. 75 fr. 350 pp.

Aptitudes and Aptitude Testing. WALTER VAN DYKE BINGHAM. Harper & Brothers, New York City, 1937. $3.00. 390 pp.

Auto-Correctivism. V. E. FISHER. Caxton Printers, Caldwell, Idaho, 1937. $3.50. 337 pp.

Biologische Typen des Menschen und ihr Verhaltnis zu Rasse und Wert. EDUARD ORTNER. Georg Thieme, Verlag, Leipzig, 1937. 7.50 M. 104 pp.

Carnegie Foundation for the Advancement of Teaching. Thirty-first Annual Report, 522 Fifth Avenue, New York City, 1936. 193 pp.

The Construction and Use of Achievement Examinations. HERBERT E. HAWKES, E. F. LINQUIST AND C. R. MANN. Houghton Mifflin Company, New York, 1936. 497 pp.

Controlled Reading. EARL A. TAYLOR. University of Chicago Press, Chicago, 1937. $3.50. 367 pp.

The Diagnosis and Treatment of Behavior Problem Children. HARRY J. BAKER AND VIRGINIA TRAPHAGEN. Macmillan Company, New York, 1936. 393 pp.

Evaluation of a Plan for Character Education. W. BRADFORD BAYLISS. Bureau of Publications, Teachers College, Columbia University, New York, 1936. $1.60. 141 pp.

Extra-Sensory Perception. J. B. RHINE. Bruce Humphries, Inc., Boston, Mass., 1935. $2.50. 169 pp.

Finding Yourself in Your Work. HARRY WALKER HEPNER. D. Appleton-Century Company, New York, 1937. $2.75. 297 pp.

Guide for Psychology. GLENN D. HIGGINSON. Macmillan Company, New York, 1937. 90 cents. 98 pp.

How to Teach Occupations. MILDRED E. LINCOLN AND JOHN M. BREWER. Ginn & Company, 1937. 134 pp.

How to Win Friends and Influence People. DALE CARNEGIE. Simon & Schuster, New York, 1936. $2.00. 237 pp.

Occupations. JOHN M. BREWER. Ginn & Company, 1936. $1.60. 622 pp.

Pitch and Intensity Characteristics of Stage Speech. MILTON COWAN. Archives of Speech, December, 1936, University of Iowa, Iowa City, Iowa. 92 pp.

Psychological Research in the Field of Radio Listening. FRANK N. STANTON. Columbia Broadcasting System, New York. 18 pp.

The Psychology of Belief. HELGE LUNDHOLM. Duke University Press, Durham, N. C., 1936. $3.00. 245 pp.

Psychology of Personality and Social Adjustment. ROBERT LEEPER. Cornell College, Mt. Vernon, Iowa. 75 cents. 61 pp.

The Reflection of Teachers' Attitudes in the Attitudes of Their Pupils. ARTHUR J. MANSKE. Bureau of Publications, Teachers College, Columbia University, New York, 1936. $1.60. 67 pp.

Report of Second Annual Conference in Guidance. Major Factors in the Problems of Youth. Department of Psychology. Fordham University Graduate School, New York, 1936. 114 pp.

Studies in General Psychology, Vol. II. CHRISTIAN A. RUCKMICK, Ed. Psychological Monographs, Vol. XLVIII, No. 2. Psychological Review Company, Princeton, N. J., 1936. 75 pp.

Student Personnel Work. E. G. WILLIAMSON AND J. G. DARLEY. McGraw-Hill Book Company, New York, 1937. $3.00. 313 pp.

Directory of American Psychological Periodicals

(Continued)

Comparative Psychology Monographs—Baltimore: Johns Hopkins University Press. Subscription $5.00. 400 pages per volume. Knight Dunlap, Managing Editor. Published without fixed dates, each number a single research. Founded 1922.

Genetic Psychology Monographs—Provincetown, Mass.: The Journal Press. Subscription $7.00 per volume. One volume per year. Edited by Carl Murchison and an international cooperating board. Quarterly. Each number one complete research. Child behavior, animal behavior and comparative psychology. Founded 1925.

Psychological Abstracts—Eno Hall, Princeton, N. J.: American Psychological Association. Subscription $6.00. 700 pages ann. Edited by W. S. Hunter and R. R. Willoughby. Monthly. Abstracts of psychological literature. Founded 1927.

Journal of Genetic Psychology—Provincetown, Mass.: The Journal Press. Subscription $14.00 per annum; $7.00 per volume. 1,000 pages annually (2 volumes). Edited by Carl Murchison and an international cooperating board. Quarterly. Child behavior, animal behavior, and comparative psychology. Founded 1821.

Journal of Social Psychology—Provincetown, Mass.: The Journal Press. Subscription $7.00. 500–600 pages annually. Edited by John Dewey, Carl Murchison, and an international cooperating board. Quarterly. Political, racial, and differential psychology. Founded 1929.

Personnel Journal—Baltimore: The Williams & Wilkins Co. Subscription $5.00. 500 pages, approximately. Edited by Walter V. Bingham. Bi-monthly.

Mental Measurement Monographs—Baltimore: The Williams & Wilkins Company. Subscription $5.00. 400 pages annually. Edited by Knight Dunlap and Buford Johnson. Published without fixed dates, each number a single research. Founded 1925.

Studies in Psychology and Psychiatry—Baltimore: The Williams & Wilkins Company. Subscription $5.00. 400 pages annually. Edited by Edward A. Paco. Published without fixed dates, each number a single research. Founded 1928.

Journal of Educational Psychology—Baltimore, Md.: Warwick & York. Subscription $6.00. Edited by Jack W. Dunlap, Harold E. Jones, and Percival M. Symonds. Published monthly except June to August. Founded 1910.

Child Development—Baltimore, Md.: Williams & Wilkins Company. Subscription $5.00. 400 pages annually. Edited by Buford Johnson. Quarterly. Founded 1930.

Character and Personality—Durham, N. C.: Duke University Press. Subscription $2.00 per annum. Quarterly. Edited by Robert Saudek (London) and Ernest Seeman (Durham, N. C.). Founded 1932.

The Psychological Exchange—11 East 125th Street, New York, N. Y. Subscription $3.00 per annum. Bi-monthly. Edited by James Hargan. Founded 1932.

Journal of Psychology—Provincetown, Mass.: The Journal Press. Subscription $14.00 per annum; $7.00 per volume. 1000–1200 pages annually. Immediate publication for any accepted manuscript. Edited by Carl Murchison and an editorial board. Quarterly. The general field of psychology. Founded 1935.

Directory of American Psychological Periodicals